



**Anton Anatolievich
Zverev**

**Modelo unificado para importação/ exportação de
dados empresariais para plataforma de dados**



**Anton Anatolievich
Zverev**

**Modelo unificado para importação/ exportação de
dados empresariais para plataforma de dados**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica do Professor Doutor Joaquim Sousa Pinto, Professor Auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro.

Dedico o meu trabalho a minha nova família, a minha esposa e a menina que esta a crescer dentro dela. É aqui que a minha vida está a começar.

o júri

presidente

Prof. Doutor Joaquim Arnaldo Carvalho Martins
Professor Catedrático da Universidade de Aveiro

vogais

Prof. Doutor Fernando Joaquim Lopes Moreira
Professor Associado de Departamento de Inovação, Ciência e Tecnologia da
Universidade Portucalense

Prof. Doutor Joaquim Manuel Henriques de Sousa Pinto
Professor Auxiliar da Universidade de Aveiro (orientador)

agradecimentos

Ao meu orientador, Professor Doutor Joaquim Sousa Pinto pela disponibilidade, motivação e todo o apoio que me deu ao longo deste trabalho demoroso. Provavelmente eu dei demasiado trabalho porque fui um dos alunos mais longos a escrever a dissertação.

A minha família que me ajudou durante estes anos todos da universidade, que me apoiaram e deram corretas direções para concluir o meu percurso universitário da melhor forma.

A minha mulher e a nossa filha que está a crescer dentro dela durante a escrita desta dissertação. Ambos me dão apoio e forças para concluir a dissertação e finalmente acabar o meu percurso universitário.

Aos meus colegas que me ajudaram ao longo de todos os anos, com os quais passei noites sem dormir para fazer algum trabalho que tínhamos de entregar as 9 de manhã do dia seguinte. Pelo convívio que tivemos. Se não fossem eles o percurso académico seria muito mais difícil e complexo.

palavras-chave

Importar conteúdo, exportar conteúdo, multilingua, Django, Python, Postgresql.

Resumo

Existem cada vez mais empresas que têm como objetivo comercializar os seus produtos no menor intervalo de tempo possível, maximizando os lucros. Para tal é necessária uma correta divulgação dos mesmos para que empresas de pequena dimensão sem capacidade de criar um site próprio possam vender os seus produtos. Um exemplo prático de um destes produtos é o portal utilizado ao longo desta dissertação como cenário em ambiente empresarial.

Como o processo de carregamento de produtos um a um para o portal, é demasiado longo e demoroso, sendo adicionalmente necessário associar as respetivas imagens e pdf's ao produto, foi encontrada a necessidade de desenvolver um processo de carregamento que permita carregar uma grande quantidade de produtos e respetivos ficheiros associados de uma só vez. Isto permite otimizar o tempo do produtor permitindo uma melhor gestão do tempo no que diz respeito ao processo de carregamento dos produtos para o portal em comparação com o processo de criação dos produtos em si. É ainda possível aplicar uma tradução multilingue para poder guardar o mesmo produto em várias línguas e dessa forma ter um mercado alvo mais alargado.

Para o desenvolvimento da aplicação foi usada a framework Django e programada na linguagem Python, guardando todo o conteúdo dos produtos em várias línguas na base de dados Postgresql.

keywords

Import/export product and file, multilanguage, Django , Python, Postgresql.

Abstract

Every time there are more and more producers and mini producers with already some quantity of products that would like to sell them in the shortest possible time to get the maximum of income and this way grow constantly. For this purpose it is necessary to disclose their products. As many small companies don't have a possibility to build their own website, they need to have other proposals like portal which was selected as a main scenario in business environment.

As a product upload for companies one by one is too long and annoying, moreover it is necessary to associate images and pdf's to each product, it was developed one general upload of many products and their files at a time to simplify a life to a producer. Besides, it is possible to apply multilanguage translations to have a possibility to save the same product in a different languages and this way get a bigger market target.

To develop this application was used Django framework and programmed in Python where all the product content was saved in different languages at Postgresql data base.

Índice

Lista de Figuras	3
Lista de Tabelas	4
Lista de Códigos.....	4
Lista de Acrónimos	5
1. Introdução	7
1.1 Enquadramento.....	7
1.2 Objetivos	8
1.3 Estrutura da dissertação	8
2. Estado da Arte	11
2.1 Alibaba.....	11
2.2 eBay.....	12
2.3 Amazon.....	14
2.4 Dropbox.....	16
2.5 Análise comparativa	16
2.5.1 Importação/exportação de produtos.....	17
2.5.2 Importação de ficheiros adicionais	17
2.6 Interação com outros sistemas	18
2.6.1 Magento	18
2.6.2 MailChimp	18
2.6.3 Google Translate API	19
3 Análise de requisitos e Arquitetura.....	21
3.1 Visão geral do sistema.....	21
3.2 Importação XLS.....	21
3.2.1 Importação Simple XLS.....	22
3.2.2 Importação Complex XLS.....	23
3.3 Arquitetura	28
3.3.1 Esquema principal do projeto	28
3.3.2 Diagramas de classes.....	28
3.3.3 Casos de uso	31

3.4 Tecnologia usada.....	32
3.4.1 Django	32
3.4.2 Python	33
3.4.3 Postgresql.....	34
4 Implementação	35
4.1 Exportação.....	35
4.1.1 Simple XLS (simples).....	36
4.1.2 Complex XLS(simples).....	39
4.1.3 Simple XLS (com dados)	41
4.1.4 Complex XLS (com dados)	42
4.2 Importação	42
4.2.1 Importação Simple XLS - Carregamento temporário	44
4.2.2 Importação Simple XLS - Confirmação	47
4.2.3 Importação Complex XLS - Carregamento temporário	49
4.2.4 Importação Complex XLS - Confirmação	52
4.2.5 Edição de produtos via XLS	53
4.2.6 Remoção de produtos via XLS	53
4.2.7 Bulk.....	53
4.2.8 Importação ficheiros	53
4.2.9 Importação ficheiro ZIP	57
4.2.10 Multilíngue	58
4.2.11 GPC categorias.....	58
4.2.12 Integração com Magento	59
4.2.13 Tratamento de erros	61
4.3 Outras importações/exportações	63
4.3.1 Importação de variáveis dinâmicas.....	64
4.3.2 Importação de categorias GPC.....	64
4.3.3 MailChimp	66
5. Resultados	69
6. Conclusão	73
6.1 Trabalhos futuros	74
6.2 Aprendizagem	74
7. Bibliografia	75

Lista de Figuras

Ilustração 1 - Esquema geral do projeto	28
Ilustração 2 - Diagrama de Classes - Tabelas portal.....	29
Ilustração 3 - Diagrama de Classes - Tabelas temporárias.....	30
Ilustração 4 - Casos de uso - Gestão de produtos.....	31
Ilustração 5 - Diagrama de atividade - Exportação XLS.....	35
Ilustração 6 - Export XLS - Fase inicial	36
Ilustração 7 - Ficheiro Simple XLS. Separador ITEM.....	37
Ilustração 8 - Ficheiro Simple XLS. Separador FILE TO	38
Ilustração 9 - Ficheiro Complex XLS. Separador ACTOR.....	39
Ilustração 10 - Ficheiro Complex XLS. Separador CATALOG TREE	39
Ilustração 11 - Ficheiro Complex XLS. Separador ITEM.....	40
Ilustração 12 - Ficheiro Complex XLS. Separador FILE TO	40
Ilustração 13 - Ficheiro Complex XLS. Separador SUPPLIERS.....	41
Ilustração 14 - Ficheiro Complex XLS. Separador PARAMETERS.....	41
Ilustração 15 - Diagrama de Actividade - Importação Simple XLS	43
Ilustração 16 - Import Simple XLS - Separador ITEM - Preenchido	45
Ilustração 17 - Import Simple XLS - Separador FILE TO - Preenchido	46
Ilustração 18 - Importação Simple XLS - Relatório	46
Ilustração 19 - Importação XLS - Apresentação de produtos.....	47
Ilustração 20 - Ficheiro Complex XLS. Separador ACTOR - Preenchido	49
Ilustração 21 - Ficheiro Complex XLS. Separador CATALOG TREE - Preenchido	50
Ilustração 22 - Ficheiro Complex XLS. Separador ITEM - Preenchido	50
Ilustração 23 - Ficheiro Complex XLS. Separador FILE TO - Preenchido.....	51
Ilustração 24 - Ficheiro Complex XLS. Separador SUPPLIERS - Preenchido	51
Ilustração 25 - Ficheiro Complex XLS. Separador PARAMETERS - Preenchido	52
Ilustração 26 - Diagrama de atividade - Importação de ficheiros.....	54
Ilustração 27 - Importação de ficheiros - Sem erros	56
Ilustração 28 - Importação de ficheiros - Com erros.....	56
Ilustração 29 - Importação XLS - Apresentação de erros.....	63
Ilustração 30 - Importação de ficheiro - Apresentação de erros	63

Lista de Tabelas

Tabela 1 - Comparação diferentes sistemas de importação de produtos	17
Tabela 2 - Comparação de diferentes sistemas de importação de ficheiros	17
Tabela 3 - Descrição detalhada do separador ITEM do SimpleXLS	22
Tabela 4 - Descrição detalhada do separador FILE TO do SimpleXLS	23
Tabela 5 - Importação Complex XLS - Validação - Separador ACTOR	24
Tabela 6 - Importação Complex XLS - Validação - Separador CATALOG TREE	25
Tabela 7 - Importação Complex XLS - Validação - Separador ITEM	26
Tabela 8 - Importação Complex XLS - Validação - Separador FILE TO	26
Tabela 9 - Importação Complex XLS - Validação - Separador SUPPLIER	27
Tabela 10 - Importação Complex XLS - Validação - Separador PARAMETERS	27
Tabela 11 - Verificações dos ficheiros do tipo imagem.....	57
Tabela 12 - Verificação dos ficheiros do tipo imagem	57
Tabela 13 - Verificações dos ficheiros do tipo pdf	57
Tabela 14 - Importação de produtos para o portal na fase de pré-carregamento.....	69
Tabela 15 - Importação de produtos para o portal na fase de confirmação	69
Tabela 16 - Diferença performance ZIP e Simple XLS	70
Tabela 17 - Comparação diferentes sistemas de importação de produtos	71
Tabela 18 - Comparação de diferentes sistemas de importação de ficheiros	72

Lista de Códigos

Código 1 - Criação da view. Hello world!	32
Código 2 - Criação do URL. Hello world!	33
Código 3 - Exemplos python.....	33
Código 4 - Python debugging	34
Código 5 - Exemplo - Execução queries.....	34
Código 6 - Ligação com API XML-RPC.....	59
Código 7 - Criação de produto em Magento.....	60
Código 8 - Criação de uma imagem em Magento.....	61
Código 9 - Execução processos em background	61
Código 10 - GPC Categorias - Estrutura XML.....	65
Código 11 - GPC Categorias - Leitura XML	65
Código 12 - Conectar-se a API MailChimp.....	66
Código 13 - MailChimp - Exportação de utilizadores - Subscrições	67
Código 14 - MailChimp - Envio da template.....	67
Código 15 - MailChimp - Envio Emails.....	68
Código 16 - MailChimp - Importação de resultados	68

Lista de Acrónimos

Acrónimo	Significado
HTTP	HyperText Transfer Protocol
XML	eXtensible Markup Language
JSON	JavaScript Object Notation
API	Application Programming Interface
PIL	Image Library
HTML	HyperText Markup Language
URL	Uniform Resource Locator
SGBD	Sistema de Gestão de Base de Dados
XLS	Ficheiros de excel
XLSX	Ficheiro de excel, formato mais compacto
GPC	GS1 standard categories
PDF	Portable document format
ZIP	Compression algorithms

1. Introdução

1.1 Enquadramento

Hoje em dia o mundo está rodeado de empresas que querem produzir e vender os seus produtos de forma rápida e eficiente. Algumas dessas empresas até compram peças produzidas por outras empresas para construir os seus produtos e os vender. Uma das partes essenciais é a produção desses produtos com uma garantia de melhor qualidade e um preço muito competitivo. No entanto, se os produtos não forem vendidos para suportar os custos da produção, então todo o trabalho realizado é desperdiçado deixando a empresa a perder dinheiro.

Para que os produtos de uma empresa sejam vendidos é necessário apostar na publicidade e em diferentes abordagens de marketing. Quanto mais eficiente for a estratégia de divulgação de um produto, maior será o alvo de pessoas que podem se interessar pelo mesmo. É possível construir o seu próprio site e disponibilizar lá os produtos, no entanto, para muitas empresas construir, manter e divulgar um site, é uma meta difícil de atingir financeiramente. Por outro lado, felizmente existem muitas maneiras de divulgar os seus produtos, muitas vezes sem gastar dinheiro. O mais comum atualmente é fazer a publicidade dos produtos através de feiras, eventos, publicidade em carros, redes sociais, lojas online, etc.

Para o responsável de uma empresa fazer publicidade e divulgação dos seus produtos, pode recorrer aos portais de publicidade ou lojas online, onde pode colocar todos os seus produtos, para que os visitantes possam ver toda a gama de produtos disponível.

Ao colocar os produtos num dos sites o responsável da empresa pode passar por etapas que podem demorar muitos dias até conseguir disponibilizar toda a informação pretendida. Além de publicar os produtos, o responsável terá que fazer associação de imagens e dos ficheiros a servir de brochura do produto. Isso tudo requer muito trabalho e muito tempo perdido a navegar de ecrã em ecrã da página para preencher todas as informações, associá-las a uma determinada categoria e ainda associar os ficheiros de cada produto. Este tempo perdido, podia ser poupado para procurar um portal diferente e fazer a mesma divulgação em outras páginas.

O público-alvo de cada produto pode ser maior se as informações do mesmo existirem em línguas diferentes, porque hoje em dia o marketing é mundial e o envio de produtos de um continente para outro é um processo habitual para quase todas as pessoas. Assim, percebemos que para melhorar a nossa divulgação é necessário que os produtos de uma empresa e a sua descrição sejam familiares para pessoas de qualquer continente. Por outro lado a tradução de todos os produtos de uma empresa para várias línguas seria um trabalho demasiado longo e cansativo.

Para simular esta situação foi desenvolvido o projeto em ambiente empresarial, que tem como objetivo divulgar todos os produtos de uma empresa para o resto do mundo.

O projeto já se encontrava na fase inicial com algumas estruturas e objetivos definidos aos quais estive que me adaptar e desenvolver a ferramenta de acordo com os requisitos pedidos. Na estrutura inicial do projeto sugeri várias alterações para garantir a melhor performance do sistema na fase de importação de produtos/imagens e a sua apresentação.

1.2 Objetivos

Um dos principais objetivos desta dissertação é apresentar uma ferramenta que consiga satisfazer todas as necessidades do utilizador perante a importação de produtos para um portal de publicidade. Da mesma forma, também é importante fazer com que a publicidade seja divulgada para utilizadores do mesmo portal.

O utilizador não pode perder demasiado tempo na inserção dos seus produtos no portal. Idealmente, este processo deverá ser feito apenas num "click" e não inserindo os produtos um a um. Todos os produtos podem ter as suas descrições, suas próprias imagens e brochuras e para isso é necessário fornecer ao utilizador uma ferramenta que consiga englobar todas estas necessidades.

A tradução de descrições é essencial para que o produto seja vendido mais rapidamente e ao mesmo tempo a sua projeção aumente. Assim, é necessário disponibilizar ao utilizador algo que seja transparente e automático e sem apresentar o mínimo de esforço.

Para além de conseguir importar todos os produtos com os seus próprios ficheiros em línguas diferentes, é também necessário fazer que estes sejam divulgados para todos os utilizadores do portal de uma forma simples e rápida. Para a divulgação é pretendido utilizar o maior comunicador do mundo que é o correio eletrónico.

Por outro lado, é importante reduzir o consumo do tempo do utilizador ao máximo e que este processo seja todo automatizado. O público alvo do portal tem de se sentir motivado ao fazer as importações de produtos e ficar satisfeito com o resultado final.

1.3 Estrutura da dissertação

A dissertação está dividida em 6 capítulos: Introdução, Estado da Arte, Análise de requisitos, Arquitetura, Resultados e Conclusões.

Introdução – neste capítulo está explicado todo o enquadramento e a necessidade de construção de uma ferramenta de importação de produtos.

Estado da Arte – neste capítulo existe um estudo de mercado de ferramentas de importação/exportação de pequenas e grandes empresas. Também é realizada uma análise e comparação da forma de importação e as suas especificações no intuito de escolher as melhores práticas para esta dissertação. São estudadas as possíveis integrações com os outros sistemas relevantes para a dissertação.

Análise de requisitos e Arquitetura – este capítulo começa pela descrição da visão geral do sistema, a seguir são descritos os tipos de importações com os respetivos requisitos. São também descritos todos os detalhes da arquitetura da ferramenta, os casos de uso e as tecnologias usadas para construção da ferramenta.

Implementação – neste capítulo são explicadas todas as implementações que foram desenvolvidas para construção da ferramenta de import/export. Por cada funcionalidade descrita são apresentados alguns resultados, tanto ao nível visual como de uma forma mais detalhada.

Resultados – neste capítulo são apresentados e analisados os resultados de alguns testes obtidos na importação.

Conclusões – neste capítulo são descritas as conclusões e os trabalhos futuros do projeto.

2. Estado da Arte

2.1 Alibaba

Alibaba[1] é um portal de uma grande dimensão que serve para vender e comprar produtos. O portal contém atualmente mais de 1 milhão de vendedores no mundo todo em que muitos deles têm milhares de produtos à venda, e constantemente os vendedores necessitam de carregar novos produtos para o portal ou atualizar os antigos com novas informações.

Para efetuar o carregamento de uma grande quantidade de produtos, Alibaba usa a ferramenta Excel que é usada na maioria dos casos para efeitos de importação de conteúdos para as bases de dados.

O ficheiro xls contém dois separadores, o primeiro separador é usado para preencher os dados do produto, no segundo separador está presente a explicação do carregamento em massa e algumas políticas do portal para poder importar corretamente os ficheiros e de acordo com as políticas do portal Alibaba.

Para começar o preenchimento de dados é necessário escolher a categoria para qual queremos importar os nossos produtos e é-nos fornecido um ficheiro xls para poder fazer o download, de seguida podemos preencher o ficheiro. O portal permite-nos, no máximo carregar ficheiros com 300 kb, isto significa, por volta de 900 produtos por cada importação. Na importação de produtos, se surgirem alguns erros, é-nos fornecido um ficheiro xls com a descrição dos mesmos. No entanto, este ficheiro não apresenta informação sobre todos os erros ocorridos. Desta forma, devemos corrigir os erros indicados e tentar novamente o upload do ficheiro com os restantes produtos. No ficheiro de erros estão especificadas as células que deram erros na primeira importação. A informação relativa e as especificações do ficheiro xls é obtida depois de um carregamento ser efetuado, no entanto, não existe documentação detalhada sobre o mesmo. [2]

A seguir a importar todos os produtos para o portal é possível fazer importação de fotografias (com limite de 3 Mb cada) e no máximo 6 ficheiros de cada vez. Este processo suporta ficheiros nos formatos jpeg, jpg e png. No final é possível ligar todos os produtos com as fotografias importadas para que as informações dos mesmos fiquem completas. [3]

Pontos positivos e negativos de Alibaba

Em relação a exportação, o Alibaba oferece uma boa abordagem para selecionar a categoria correta. As categorias são divididas por 3 grupos, sendo a primeira geral e as outras subcategorias. Desta forma é fácil conseguir escolher a categoria adequada aos produtos.

Alibaba também fornece uma ferramenta relativamente simples de usar. Tem poucos campos para preencher sobre um produto, não baralhando tanto o utilizador. Mais um ponto positivo é a ajuda que é dada com as dropdowns, uma vez que na maioria dos casos os produtos que vão ser importados tem campos com uma lista de valores para escolher/preencher. Desta forma, se por alguma razão o utilizador se enganar a escrever, esse campo será mais um erro que ele terá de observar e corrigir. No entanto, por vezes, a correção dos erros apresenta algumas dificuldades. A

descrição do funcionamento do preenchimento do xls também é importante, uma vez que o utilizador pode facilmente, com a menor dúvida e sem ter que mudar de aplicação, saber todas as respostas.

O tamanho dos ficheiros que o Alibaba permite importar é, no entanto, muito pequeno. Com pouca descrição de produtos o ficheiro xls facilmente ultrapassa os 300 kb. Se os produtos inseridos no ficheiro xls apresentarem uma descrição demasiado grande, então o número máximo de produtos a transferir será relativamente reduzido e assim alguns dos utilizadores necessitam de criar vários ficheiros para poder importa os seus produtos para o portal Alibaba.

O tratamento de erros também devia ser um pouco diferente. Não é nada pratico fazer download do ficheiro xls que indica os erros cometidos pelo utilizador, obrigando desta forma o utilizador a ter vários ficheiros xls abertos para conseguir terminar a importação com sucesso.

A impossibilidade de importar ficheiros .xlsx também é um ponto negativo dado que a maioria dos ficheiros atualmente gerados pelos utilizadores já vêm com a extensão nova. Por fim, o ficheiro xls com a descrição dos campos de importação devia ser disponibilizado em mais que uma língua para além do inglês.

No sistema Alibaba também só é possível transferir três tipos de ficheiros, não necessitam de mais nenhum tipo para associar aos produtos, e isso está bem explicado ao importar. Ao importar os ficheiros para o sistema é dado um feedback geral da importação, apresentado com uma barra.

Um dos problemas de importação desses ficheiros é ser 6 de cada vez, mesmo que cada produto no máximo consegue receber 6 ficheiros para associar. Isto é não é prático, porque sendo assim para 100 produtos é necessário fazer este tipo de importação no mínimo 100 vezes.

O tamanho máximo a importar penso que seja bom porque uma imagem tirada com uma máquina fotográfica com relativamente boa qualidade não ultrapassara esse limite.

A validação de ficheiros também é importante para que não sejam inseridos ficheiros errados para o portal. E quando os ficheiros não são válidos e têm extensão diferente da permitida é dado um feedback relativamente viável para compreender a razão do erro.

No geral consegui tirar algumas ideias importantes do Alibaba para pensar em boas soluções para a ferramenta que vou desenvolver.

2.2 eBay

eBay[4] é atualmente o maior portal do mundo de venda, compra e leilão de bens. Neste momento, o eBay contém mais de 181 milhões de utilizadores que podem vender e comprar os seus produtos.

Para disponibilizar os seus produtos no sistema é necessário preencher o formulário do produto e carregar as fotografias, mas se existir uma enorme quantidade de produtos a inserir já é necessário utilizar as ferramentas disponibilizadas pelo eBay, entre quais estão ficheiros Excel e

ficheiros csv. Como os ficheiros csv podem ser criados através do ficheiro Excel a utilização dos ambos os casos é semelhante.

O eBay disponibiliza vários tipos de ficheiros excel, para inserção de produtos e categorias. Para começar a introduzir os seus produtos no ficheiro é necessário descarregar o ficheiro e preenchê-lo conforme as instruções indiquem. [5]

No eBay é possível efetuar importações para a plataforma de no mínimo 50 produtos e no máximo 5000. Depois de importar os produtos estes aparecem na plataforma eBay e a partir daí é possível importar e associar imagens aos produtos.

Pontos positivos e negativos de Ebay

Na exportação do eBay não existe nada que diferencie para que tipo de categoria vão ser importados os produtos, o único ficheiro que existe que possa fazer isso é o *CatalogTemplate*, mas não há nada nesse ficheiro que ajude a diferenciar as categorias.

O ficheiro xls oferecido pelo eBay é relativamente fácil de preencher, no entanto complexo, dada a quantidade de campos existentes para preencher. A ajuda sobre cada coluna é feita em comentários, que facilita muito o processo de preenchimento do Excel. No entanto, quando já temos vários produtos preenchidos, o aparecimento dos comentários pode tornar-se um pouco chato para o utilizador.

A possibilidade de adicionar imagens ao produto é benéfica. No entanto, adicionar só uma imagem e sempre através de URL acaba por não ser uma boa solução. Dado que um produto normalmente é representado no mínimo, por 2 ou 3 imagens, para inserir as imagens, é necessário encontrar uma página web onde estas possam ser alojadas e depois disponibilizar o URL na importação. Este processo, para muitos utilizadores inexperientes, poderá ser demasiado complicado.

O eBay permite importar entre 50 a 5000 produtos por cada ficheiro xls. O limite máximo é suficiente porque raramente uma empresa tem mais que 5000 produtos e mesmo se existir uma grande quantidade de produtos não é bom gerir tudo no mesmo xls, porque torna-se muito pesado e complicado. No entanto, muitas empresas não possuem 50 produtos, e como foi explicado nas páginas anteriores, importar produto a produto, é uma tarefa longa e chata. Assim, para evitar processos de importação demasiado trabalhosos, o limite mínimo de importação, nunca deveria passar dos 5 produtos.

O eBay exige que a importação do ficheiro xls seja feita através do ficheiro .txt, para isso é necessário converter o ficheiro para outro tipo e efetuar a importação e sempre que o ficheiro tiver erros de inserção é necessário abrir o ficheiro xls, alterá-lo novamente, guardá-lo em formato .xls e importar novamente. Apesar de não ser uma tarefa complicada, é sempre um passo extra para o utilizador. Da mesma forma, para alguns utilizadores, pode mesmo ser um quebra-cabeças, porque tem pouca experiência em trabalhar com as aplicações do Windows/Linux/Mac.

No geral não consegui tirar muitos pontos positivos do processo de importação do eBay mas os pontos negativos levaram-me a tirar as minhas conclusões para conseguir construir algo que não apresente estes erros.

2.3 Amazon

Amazon[6] é um portal multinacional para venda de produtos novos ou usados. Foi uma das primeiras empresas a realizar as vendas de produtos pela internet. Os produtos podem ser vendidos por utilizadores particulares ou mesmo pequenas, médias e grandes empresas.

Quando falamos em vendas dos produtos por parte particular não surge nenhum problema uma vez que a quantidade de produtos que essa pessoa normalmente quer vender não é muito grande. No entanto, o problema pode surgir quando as empresas com uma quantidade de produtos relativamente grande querem vendê-los e para isso necessitam de os publicar no portal da Amazon. Naturalmente, carregar os produtos um a um é uma perda de tempo para estes utilizadores que usam a plataforma Amazon.

Para os seus carregamentos de dados para empresas com uma grande dimensão a Amazon fornece duas soluções.

Para a primeira solução podem ser usados os ficheiros .txt, .csv e .xls. Os ficheiros deste tipo só são precisos para preenchimento de informação relativa aos produtos. Assim, para prosseguir com a importação é necessário converter os ficheiros já construídos para formato .txt [7].

Os produtos importados para o portal podem pertencer a categorias diferentes. Desta forma, para a escolha de uma categoria a Amazon oferece várias ligações que correspondem a categorias diferentes [8]. Os ficheiros descarregados são de diferentes tipos, desde de categorias tradicionais até ao upload de música, livros e filmes. Todos os ficheiros transferidos tem o formato .xls. Para a transferência de imagens para os produtos é necessário indicar o URL de cada imagem.

Se na importação de produtos para o portal surgirem erros, estes serão apresentados como um determinado tipo de erro. Os detalhes dos erros poderão depois ser consultados num ficheiro pdf a parte [9].

A segunda solução oferecida pela Amazon é o uso das macros do Excel [10], no entanto, a explicação do uso destas macros só existe para o Excel 2007. Quando as macros tiverem ativas é possível preencher o ficheiro xls e efetuar logo a importação de produtos novos ou a alteração de produtos já existentes no sistema. Para utilizar esta solução é requisito ter Microsoft Excel 2002, 2003 ou 2007.

A indicação de erros durante a importação é dada diretamente no Excel com uma *dialogbox*, e o ficheiro com os seus produtos não será transferido até que todos os erros forem resolvidos [11].

A importação de produtos para a Amazon só pode ser efetuada com uma conta de vendedor (*Seller Account*). A importação através de Excel não pode ultrapassar os 50000 produtos e o

tamanho do ficheiro xls não pode ser superior a 60 Mb [12]. Depois de importar os produtos, estes permanecem em espera durante 24 horas para serem aprovados pela Amazon.

Pontos positivos e negativos de Amazon

A exportação dos ficheiros xls é feita pelas categorias, fornecendo uma lista de links para diferentes tipos. Para algumas categorias existe mais do que um link, o que pode confundir o utilizador do portal. O ficheiro exportado ocupa por volta de 6 Mb por causa das suas explicações com imagens dentro do ficheiro. Um ficheiro com esse tamanho acaba por ser acima do normal, tornando desta forma o ficheiro ser pesado e mais lento em alguns computadores de baixa performance.

Ao preencher o ficheiro xls da Amazon existe bastante ajuda, entre 6 separadores do ficheiro 5 deles são de ajuda. Explicam que características pode ter cada campo, os exemplos de preenchimento, os valores que são válidos em determinados campos e a descrição como deve ser as imagens de cada produto. No entanto, para a quantidade de campos apresentada, ainda falta alguma ajuda. As dropdowns, são muito importantes para não estar constantemente a mudar de separador para ver o que é necessário escrever em determinada célula.

O limite de produtos a inserir é grande, o que é uma parte importante da ferramenta. Muitas das empresas só tem 10 produtos e não querem ter trabalho de inserir produtos um a um e também pelo contrário há empresas que tem 10000 produtos e querem os inserir de uma só vez. A Amazon permite este tipo de importações.

Inserção de produtos através de macros é um outro ponto que pode ser considerado positivo, mas é bastante complicado para alguns utilizadores. A inserção deste tipo não necessita de efetuar o upload de ficheiro xls no portal, mas sim fazer a importação diretamente através do Excel. Os erros no xls serão apresentados logo, sem sequer fechar o ficheiro. Para este tipo de importações é necessário configurar no Excel as macros que farão a importação pelo utilizador. A configuração das macros para alguns dos utilizadores pode ser muito complexa. Nem todos os utilizadores pretendem constantemente inserir produtos no portal, mas sim inserir uma só vez, por isso a configuração das macros seria uma parte onde se perde bastante tempo. A existência desta ferramenta pode ser útil para empresas que estão constantemente a efetuar alterações sobre os produtos e que fazem a adição de produtos novos com regularidade.

A importação das imagens juntamente com os produtos é um ponto positivo, mas temos a mesma situação como no eBay. Podem ser inseridas até 10 imagens por produto, mas todas elas tem de ser inseridas através do URL da imagem. Essa abordagem pode trazer aos utilizadores problemas, como por exemplo, o utilizador necessita de ter algum repositório para poder guardar as suas imagens e depois copiar o URL das mesmas e inserir no xls. Para alguns utilizadores até copiar um simples URL da imagem pode ser complicado.

A importação para o portal Amazon tem de ser feita através do ficheiro txt, como no eBay, neste caso temos novamente o problema de utilizadores com pouco conhecimento de diversas ferramentas não conseguirem efetuar esta operação com facilidade, mesmo com a documentação suficientemente boa.

No geral existe alguns pontos positivos no portal Amazon, mesmo sendo com alguma complexidade. Mas é possível tirar as suas conclusões e aplicar esses pontos na ferramenta que estamos a desenvolver e até tornar os pontos positivos um pouco mais amigáveis para os utilizadores, principalmente para aqueles que estão pouco familiarizados com os sistemas operativos e com as ferramentas existentes no mundo informático para realizar certas operações.

2.4 Dropbox

Dropbox [13] é um portal que serve de repositório de ficheiros de todo o tipo. O principal objetivo desta plataforma é inserir todos os ficheiros que o utilizador pretenda guardar numa *cloud*.

Os ficheiros para a importação que a dropbox aceita tem o limite máximo de 10 Gb e não existe qualquer tipo de verificação se o ficheiro importado é válido ou de confiança. [14]

A importação de ficheiros para a dropbox pode ser feita a partir da página web, selecionando os ficheiros pretendidos. Assim, todos os ficheiros serão postos numa fila e é dado um feedback de upload com sucesso de cada ficheiro. Os ficheiros também podem ser importados a partir do computador, previamente instalando a aplicação da Dropbox. Por fim, é também possível carregar os ficheiros a partir do telemóvel. Sempre que os ficheiros são importados é dado um feedback para o utilizador perceber o estado da transferência.

Pontos positivos e negativos de Amazon

Na Dropbox é possível transferir qualquer tipo de ficheiros sem alguma preocupação, pois é um sistema de repositório de ficheiros. O tamanho máximo permitido também é muito grande e é transferido desde que a conta permita. Para importar os ficheiros não existe qualquer limite do número máximo de transferências.

Ao importar os ficheiros é dado um feedback sobre cada um, que só é importante para ficheiros de grande dimensão, mas para ficheiros pequenos não seria tão necessário. A validade dos ficheiros não é efetuada, fica totalmente da responsabilidade do utilizador que ficheiros estão a ser transferidos. Por isso os erros a apresentar só existem caso houver alguma falha na transferência de ficheiros.

2.5 Análise comparativa

Ao estudar todas as tecnologias já existentes tirei alguns exemplos que vão servir como solução para o meu projeto. É necessário ter em conta todas as tecnologias, porque uma pode completar a outra. Da mesma forma, serão ainda adicionadas funcionalidades novas que, do meu ponto de vista, são essenciais para o sucesso do portal. Por outro lado, nem todas as funcionalidades existentes são importantes, porque muitas das vezes podem baralhar mais o utilizador final.

2.5.1 Importação/exportação de produtos

A parte de importação de produtos para o portal é a principal preocupação. Para isso foram estudados os sistemas como Alibaba, eBay e Amazon, todos estes sistemas tem como objetivo principal publicar e vender os produtos dos seus clientes. Todos os sistemas fornecem as suas ferramentas de importação de produtos das quais, do meu ponto de vista, há pontos positivos e negativos. A Tabela 1 mostra uma breve descrição sobre eles.

Descrição\Sistemas	Alibaba	eBay	Amazon
Tipo de importação	Xls	txt	txt e macro
Quantidade de produtos a importar	+ 900 produtos	50-5000 produtos	50000 produtos
Descrição do excel	Sim	Sim	Sim
Dropdown	Sim	Não	Não
Associar imagens	Não	Sim, uma por link	Sim, por link
Detalhe de erros	No novo excel	Lista de erros	Lista de erros e no excel
Tamanho máximo do ficheiro	300 Kb	Não tem	60 Mb
Facilidade de importação (1-difícil até 4-fácil)	3	2	2
Exportação xls vazio	Árvore de categorias	Sem categorias	Diferentes downloads do xls
Exportação com dados	Não	Não	Não

Tabela 1 - Comparação diferentes sistemas de importação de produtos

2.5.2 Importação de ficheiros adicionais

A importação de ficheiros adicionais é um extra muito significativo para cada produto, porque a imagem de um produto ou a sua brochura podem influenciar positivamente o cliente a comprar o produto. Para isso estudei o sistema de importação do Alibaba e Dropbox e tirei algumas ideias para poder evoluir este tipo de transferência de ficheiros. Na tabela em baixo são apresentados os pontos negativos e positivos dos sistemas estudados.

Descrição\Sistema	Alibaba	Dropbox
Tipo de ficheiros	jpg, jpeg, png	Todos
Tamanho máximo	3Mb	10Gb
Nº de ficheiros a importar	6	Sem limite
Upload feedback	Geral	Em cada ficheiro
Apresentação do erro	Sim, com detalhe	Sim
Validação da extensão	Sim	Não
Validação do ficheiro	Sim	Não

Tabela 2 - Comparação de diferentes sistemas de importação de ficheiros

2.6 Interação com outros sistemas

No mundo real quase todos os sistemas precisam de usar ferramentas já existentes para facilitar o trabalho e diminuir tempo em implementação de uma determinada tarefa. Muitas das vezes começamos a desenvolver algo e não perdemos tempo a pesquisar, mas isso nem sempre é um bom caminho, por isso foram estudadas algumas ferramentas para realizar algumas tarefas que o nosso portal necessita, de forma transparente e rápida.

2.6.1 Magento

Magento [15] é uma plataforma de venda na internet que é opensource e foi desenvolvida em PHP e MySQL. Esta plataforma pode ser adaptada a qualquer estilo e efetuar quaisquer operações, visto que foi desenvolvida para ser totalmente modular. Hoje em dia, esta plataforma é utilizada por muitas empresas para vender os seus produtos na internet.

A plataforma Magento foi comprada pelo eBay em 2011 e tem crescido cada vez mais em termos de utilização por empresas de venda. Este ano, em 2015, foi considerada a melhor [16] plataforma de eCommerce[17].

A integração com o Magento é essencial para publicar os produtos numa loja online quando estamos a carrega-los para o catálogo. Para evitar o trabalho de criar uma loja online própria visto que o Magento já nós fornece uma serie de funcionalidades como pagamentos de produtos via paypal ou cartão de credito.

Sem integrar a importação diretamente com o Magento era necessário criar no seu próprio portal e as funcionalidades que o Magento traz, ou se usar o Magento ter que carregar todos os produtos novamente.

2.6.2 MailChimp

MailChimp [18] é uma plataforma que permite enviar as newsletters para os utilizadores. Esta plataforma foi desenvolvida para facilitar muito o trabalho das pessoas que pretendem enviar grande quantidade de emails para diferentes utilizadores do seu portal ou simplesmente utilizadores que queiram subscrever a uma newsletter. Além de enviar apenas emails simples existe a possibilidade de personalizar os emails de acordo com os dados dos utilizadores, como por exemplo com o nome do cliente.

Sem esta plataforma seria muito complicado informar os utilizadores, de forma personalizada, sobre uma nova ferramenta do sistema. E também informar os utilizadores do sistema de existência de novos produtos no portal. A ferramenta MailChimp trata de uma forma transparente e simples este processo. Isto permite-nos poupar imenso tempo para construir um mail diferente para cada um dos clientes com uma lista de novos produtos inseridos.

2.6.3 Google Translate API

A Google Translate é uma plataforma de traduções de um texto de uma língua para outra qualquer à escolha. Esta ferramenta tem o seu próprio sistema de traduções [19], utilizando machine learning que ao longo do tempo "aprende" a maneira correta de traduzir determinadas palavras ou frases.

A tradução de informação sobre os produtos é uma parte muito essencial para o portal, visto que os produtos podem ser vistos por diferentes partes do mundo e para isso é necessário que um produto esteja exposto em várias línguas.

A Google Translate API [20] permite nos efetuar traduções instantâneas usando a ferramenta de traduções da Google. Visto que o portal é de multilíngue o utilizador pode não saber escrever numa língua diferente da nativa e pode optar por opção de traduzir logo os produtos que estão a ser inseridos, corrigindo futuramente as traduções caso seja necessário.

Isto permite evitar de carregar os mesmos produtos várias vezes. Da mesma forma é efetuado logo a tradução de todos os produtos para as línguas que o utilizador escolher.

3 Análise de requisitos e Arquitetura

3.1 Visão geral do sistema

Depois de uma análise de todos os sistemas existentes em relação a importação e exportação de produtos decidiu-se seguir a mesma lógica que os sistemas atuais usam. Para além de usar algo parecido pensei em avançar para além daquilo que os outros sistemas oferecem.

A tradução de produtos para todas as línguas que o utilizador pretender é uma inovação neste tipo de processo. Pode não significar muito para alguns utilizadores mas de certeza que é uma grande ajuda a alguns deles em ter os seus produtos em línguas distintas, atraindo desta forma o maior alvo de pessoas.

Importação e associação de ficheiros de tipo imagem e pdf também é uma parte essencial para desenvolver no projeto. Por vezes os utilizadores necessitam de transmitir aos clientes mais informação do que simplesmente em imagens, o que os sistemas existentes não fazem.

Durante a importação de todo o tipo de ficheiros o relatório com os erros é essencial. Da mesma forma, quanto mais detalhado for melhor para o utilizador facilmente perceber o que é, onde é, e como pode corrigir e efetuar novamente a importação com sucesso. Alguns dos sistemas estudados não tem a explicação do erro ao pormenor.

A importação e o tratamento de todos os ficheiros deverá ser desenvolvida num sistema que permitira efetuar várias abordagens sobre os ficheiros para conseguir escolher a mais eficiente. Assim, conseguimos garantir aos utilizadores a melhor apresentação e divulgação do seu produto.

O relatório deverá ser apresentado com todos os detalhes que o utilizador introduzira durante a importação e com possibilidade de verificar os produtos importados para garantir a introdução correta. Este tipo de relatórios pode ser apresentado com o código HTML que consegue criar e transmitir ao utilizador o aspeto desejado.

3.2 Importação XLS

Aqui serão apresentados os principais requisitos da importação de produtos para o portal. Estes requisitos são relativamente aos dois tipos de XLS. O primeiro tipo é um XLS simples que tem a mínima complexidade que serve para ser usado pelos utilizadores com pouca experiência, fornecendo só a possibilidade de importar os produtos e os ficheiros associados. O segundo tipo é um ficheiro XLS complexo que é usado por utilizadores mais avançados, que requer um bom conhecimento sobre o sistema. Através do XLS complexo é possível inserir no sistema, além dos produtos e ficheiros associados, os dados da empresa, o seu catalogo para os produtos e os fornecedores de cada produto.

3.2.1 Importação Simple XLS

Para manipular os produtos de forma simples é necessário criar um XLS mais simples que contem só 3 separadores em que o ultimo é uma explicação técnica de procedimentos de preenchimento do ficheiro, para que este seja importado sem erros.

Cada um dos campos do ficheiro XLS tem que ter as suas especificações (Tabela 3) que são verificadas durante o processamento na primeira fase. No caso de sucesso serão produzidos resultados para verificar os produtos inseridos no ficheiro pelo utilizador e caso esteja tudo correto é possível continuar com o processamento de inserção dos produtos no portal. Os campos dinâmicos têm de corresponder à lista que está inserida na célula.

No caso do insucesso deverá ser gerado um relatório de erros que será explicado mais a frente em ponto 4.2.13 Tratamento de erros.

Nome	Detalhes				
	Obrigatório	Min.(Char)	Max.(Char)	Exemplo	Dropdown
Nome do produto	SIM	2	50	Produto mel	-
Características	SIM	1	2000	Características	-
Keyword	SIM	1	100	Mel	-
Código de barras	NÃO	-	50	5014328028740	-
Ref. Interna	SIM	1	18	Ref. 0001	-
Atributos	NÃO	-	-	-	Lista GPC
Marca	NÃO	-	100	Auchan	-
Modelo	NÃO	-	100	Verão	-
Especificações	NÃO	-	2000	Intenso	-
Língua	SIM	2	2	pt	-

Tabela 3 - Descrição detalhada do separador ITEM do SimpleXLS

Cada linha do xls corresponde a um produto que ficará no portal com informação que vai ser preenchida.

No separador FILE TO cada linha identifica um ficheiro, que vai ser associado a um produto. Para que a importação seja corretamente inserida, o utilizador deve, durante o preenchimento, respeitar a informação que esta detalhada na Tabela 4 - Descrição detalhada do separador FILE TO do SimpleXLS.

Só é possível associar 4 imagens por produto, esta limitação existe para não sobcarregar a página dos produtos no portal. O número de pdf's não esta definido.

Nome	Detalhes					
	Obrigatório	Min.(Char)	Max.(Char)	Exemplo	Dropdown	Requisitos
Tipo	SIM	-	-	IMG	IMG, PDF	
Nome do ficheiro	SIM	1	40	ProdMel	-	
Ref. Interna	SIM	1	18	Ref. 0001	-	Existir

Ordem	SIM	1 número	1 número	1	-
Alt	NÃO	-	30		-
Tooltip	NÃO	-	50		-

Tabela 4 - Descrição detalhada do separador FILE TO do SimpleXLS

Da mesma forma, como foi explicado, para o primeiro separador existe um tratamento de erros que produz um relatório que será explicado no ponto 4.2.13 Tratamento de erros.

3.2.2 Importação Complex XLS

A importação do Complex XLS também deverá consistir por duas partes, sendo a primeira a fase de pré-carregamento, necessária para a verificação de erros do xls e a sua verificação por parte do utilizador. A segunda parte consiste na confirmação dos produtos para estes serem importados para o portal.

O ficheiro Complex XLS deverá conter 7 separadores, em quais é possível inserir toda a informação sobre a empresa, produtos, fornecedores, ficheiros a associar, etc. Para que a importação esteja correta é necessário respeitar as regras de validação.

No caso do insucesso deverá ser gerado um relatório de erros que será explicado mais a frente em ponto 4.2.13 Tratamento de erros.

O primeiro separador contém os dados da empresa(ver Tabela 5) que contém campos obrigatórios para garantir os campos mínimos pedidos pelo portal. Este separador contém só uma linha de registo, ignorando as linhas seguintes caso estas estejam preenchidas.

Para além destas regras é necessário verificar:

- Utilizador tem de existir no portal;
- NIF da empresa nova é um campo único e não pode existir no portal, caso exista, o utilizador é informado que a empresa com o NIF inserido já existe no portal;
- Email é mais um campo único e é apresentado um erro caso o email inserido exista no sistema.

Nome	Detalhes				
	Obrig.	Min.(Char)	Max.(Char)	Exemplo	Requisitos
Nome do utilizador	SIM	1	30	Antonza	-
Banner cor	SIM	7	7	#B0B0B0	Válido
Nome da empresa	SIM	1	255	Antonza Comp	-
Nome curto empresa	SIM	1	30	Anton Co	-
Cod. Pais	SIM	2	2	PT	Válido
NIF	NÃO	-	32	123456789	Válido
Morada 1	SIM	1	255	Aveiro	Sem # e &
Morada 2	NÃO	-	255	Rua da Pé	Sem # e &
Morada 3	NÃO	-	255	-	Sem # e &

Morada 4	NÃO	-	255	-	Sem # e &
Código postal	SIM	1	9	3800-000	Válido por país
GPS coordenadas	NÃO	-	40	N 40.6, E -8.6	Válido
Indicativo país	SIM	-	-	00351	Número válido
Telemóvel	NÃO	-	25	919191919	Número
Telefone	NÃO	-	25	282126594	Número
Fax	NÃO	-	25	282123654	Número
Mail	SIM	1	75	antonza@ua.pt	Válido
URL	NÃO	-	255	www.anton.pt	Válido

Tabela 5 - Importação Complex XLS - Validação - Separador ACTOR

O segundo separador contém a informação do catálogo da empresa(ver Tabela 6) que está registada linha a linha e em que cada linha representa o nome do catálogo/sub-catálogo.

Caso seja encontrado pelo menos um erro numa das linhas então a importação não será possível, devido a inconsistência da árvore do catálogo. O resto das linhas do catálogo serão avaliadas na mesma e no caso de existir mais erros serão apresentados ao utilizador.

Cada linha do separador corresponde a uma linha do catálogo. Se existir uma linha sem qualquer preenchimento então será ignorada, continuando a verificação das linhas seguintes.

No entanto, para além destas regras, é necessário que mais alguns casos em concreto sejam respeitados para garantir a correta estrutura do catálogo:

- Tem de existir sempre pelo menos um nível de catálogo e esse nível tem de estar na primeira coluna;
- O campo de dependência do nível 1 do catálogo tem de ter valor 0, indicando desta forma que é a estrutura pai do catálogo;
- O número do catálogo é o campo único e por isso não pode haver repetições do mesmo número do catálogo;
- A dependência tem de ser um número de um catálogo que esteja inserido nas linhas anteriores;
- Caso alguma das linhas não seja preenchida ela é ignorada e será lida uma próxima linha que tenha informação;
- A estrutura do catálogo não pode passar do nível 1 para nível 3, 4, etc, tem de seguir a sequência.

Nome	Detalhes				
	Obrig.	Min.(Char)	Max.(Char)	Exemplo	Requisitos
Nível 1	SIM	1	100	Sobre	-
Nível 2	NÃO	-	100	Nós	-
Nível 3	NÃO	-	100	Nível 3	-
Nível 4	NÃO	-	100	Nível 4	-
Nível 5	NÃO	-	100	Nível 5	-

Nível 6	NÃO	-	100	Nível 6	-
Número do catalogo	SIM	-	-	0	Existir nº 0
Dependência	NÃO	-	-	1	Existir nº do catalogo
Nome do catalogo	SIM	2	35	Produto	Sem --, // e %
Título do catalogo	NÃO	-	100	Título	-
Info do catalogo	SIM	-	-	3800-000	-

Tabela 6 - Importação Complex XLS - Validação - Separador CATALOG TREE

O terceiro separador contém a informação de produtos(ver Tabela 7) que vão ser inseridos no portal. Esta informação tem de ser respeitada pelo utilizador, caso contrario a linha do registo será ignorada e apresentada com respetivos erros ao utilizador. Cada linha do separador corresponde a um produto

No entanto, além destas regras é necessário garantir alguns pontos importantes para o preenchimento correto dos produtos:

- A primeira coluna contém a operação que irá ser realizada sobre o produto.
 - IGNORE - Ignora totalmente a linha do separador;
 - ADD - Adição de um novo produto no portal. Tem de ser preenchido pelo menos um dos campos da referência interna;
 - MODIFY - Modificar um produto existente no portal. Tem de ser preenchidos obrigatoriamente referência interna e nova;
 - DELETE - Apagar um produto. A remoção do produto é feita através da referência interna do produto, caso o produto não exista é registado um erro;
- Todas as referências internas têm de ser únicas, tanto no separador como no portal;
- Papel de cada produto é construído por 6 tipos e cada produto tem de ter obrigatoriamente pelo menos um tipo. Para registar diferentes tipos de papel por produto é necessário inserir os números do role separados por virgula, ponto ou espaço. Ao ler esta célula é feito um *split* por qualquer símbolo indicado e desta forma registados os papéis do produto.
- A coluna dependência tem de ser preenchida com o número do catálogo a qual vai pertencer um produto. O produto pode estar associado a mais do que um catálogo, neste caso os números de catálogos são separados por vírgula, ponto ou espaço.

Nome	Detalhes				
	Obrig	Min.(Char)	Max.(Char)	Exemplo	Requisitos
Ação	SIM	-	-	ADD	ADD, MODIFY, DELETE, IGNORE
Código barras	NÃO	-	36	1212541	Válido
Ref. interna	SIM	1	18	Ref.001	Única
Nova ref. interna	NÃO	-	18	Ref.002	Ação = MODIFY
Nome	SIM	-	50	Produto 1	-
Características	SIM	1	2000	Caracter.	-
Especificações	NÃO	-	2000	Spec	-

Papel	SIM	-	6	1,2	Existir correspondência
Dependência	SIM	-	-	3	Existir no CATALOG TREE
Ref. Fabricante	NÃO	-	18	RefFab.01	-
NIF fabricante	NÃO	-	32	32165987	Válido
Nome fabricante	NÃO	-	255	Aveiro	-
URL fabricante	NÃO	-	255	ww.ua.pt	Válido

Tabela 7 - Importação Complex XLS - Validação - Separador ITEM

O quarto separador contém a informação de todos os ficheiros que serão associados a produtos, linhas de catálogo ou o logotipo de empresa. Cada linha deste separador corresponde a um ficheiro

Para que o ficheiro seja associado corretamente, os registos deste separador tem de respeitar as regras da Tabela 8. Para além das regras da tabela é necessário respeitar as seguintes regras de integração:

- A primeira coluna é uma lista de valores predefinidos que indica qual é o destino do ficheiro a associar:
 - ACTOR - ficheiros associados a empresa do utilizador. Este registo é único e só pode ser de tipo imagem. Nota: o ficheiro na fase de transferência tem de ser do tipo .png;
 - CATALOG - ficheiros associados a linha de catálogo. A cada linha de catálogo só pode ser associada uma imagem e pdf's.
 - ITEM - ficheiros associados aos produtos da empresa. Cada produto pode ter no máximo 4 imagens e pdf's.
- A segunda coluna do separador indica o tipo de ficheiro. Só existe dois tipos de ficheiros, IMG - imagem e PDF - pdf. As características e detalhes dos ficheiros encontram-se no 4.2.8 Importação ficheiros.
- Na coluna de destino é necessário indicar o número do catálogo, referência do produto ou shortName da empresa para que determinado ficheiro fique associado ao destino correto. Caso o destinatário não existe é apontado como erro;

Na coluna nº ordem é indicada a ordem como os ficheiros irão aparecer no portal. No caso do catálogo é a ordem dos pdf's e no caso do produto é a ordem das 4 imagens.

Nome	Detalhes				
	Obrig.	Min.(Char)	Max.(Char)	Exemplo	Requisitos
Para	SIM	1	100	ITEM	ACTOR, ITEM, CATALOG
Tipo	SIM	-	-	IMG	IMG ou PDF
Nome do ficheiro	SIM	1	40	FileName	Special char
De	SIM	-	-	Ref.001	Nº cat, ref. interna, actor
Ordem	SIM	-	-	1	Número
Alt	NÃO	-	30	Alt	-
Tooltip	NÃO	-	50	Tooltip	-

Tabela 8 - Importação Complex XLS - Validação - Separador FILE TO

No quinto separador contém toda a informação sobre o fornecedor dos produtos. Cada linha do separador corresponde a um fornecedor de um produto. Cada produto pode ter mais que um fornecedor e um fornecedor pode ter vários produtos.

Cada linha deste separador será verificada por algumas regras da Tabela 9. Para além destas regras, ao preencher é necessário validar que a referência interna do produto existe no separador ITEM. Caso seja o mesmo fornecedor para vários produtos é necessário de preencher uma linha por cada produto do fornecedor.

Nome	Detalhes				
	Obrig.	Min.(Char)	Max.(Char)	Exemplo	Requisitos
Ref. interna	SIM	-	100	Ref.001	Existir no ITEM
Ref. fornecedor	NÃO	-	-	RefF.01	-
NIF do fornecedor	NÃO	-	32	123564	Válido
Nome do fornecedor	SIM	-	255	Cola	-

Tabela 9 - Importação Complex XLS - Validação - Separador SUPPLIER

No sexto separador só existe uma linha e duas colunas, que correspondem a parametrização do ficheiros complex XLS. O resto das linhas para além da primeira serão ignoradas, visto que não é possível indicar mais que um tipo de parâmetros no ficheiro.

Os dados registados neste separador têm de passar pelas regras descritas na Tabela 10. A primeira coluna indica a língua em que a empresa, o catálogo e os produtos vão ser importados. A segunda coluna só existe para modificações ou remoção dos produtos do portal. As ações estão identificadas na primeira coluna do separador ITEM.

Nome	Detalhes				
	Obrig.	Min.(Char)	Max.(Char)	Exemplo	Requisitos
Língua	SIM	2	2	PT	Válido
ID do actor	NÃO	-	-	1245	Existente no portal

Tabela 10 - Importação Complex XLS - Validação - Separador PARAMETERS

3.3 Arquitetura

3.3.1 Esquema principal do projeto

O projeto está dividido em várias partes (Ilustração 1) para facilitar a sua organização e dividir o trabalho de uma forma correta entre a equipa.

No projeto existe duas partes lógicas para garantir o funcionamento correto do portal, que são FrontEnd e BackEnd. FrontEnd corresponde à parte visual do portal, que garante a melhor apresentação de informação que está armazenada nas bases de dados do portal. Para aceder a bases de dados e ao repositório de imagens e pdf's é utilizada a estrutura do BackEnd que efetua os pedidos vindos do FrontEnd na base de dados e repositório armazenados no AWS Amazon e devolve a informação necessária ao FrontEnd.

O BackEnd é responsável pela comunicação do utilizador com as bases de dados e através do Backend é que são processados todos os ficheiros que o utilizador pretende importar.

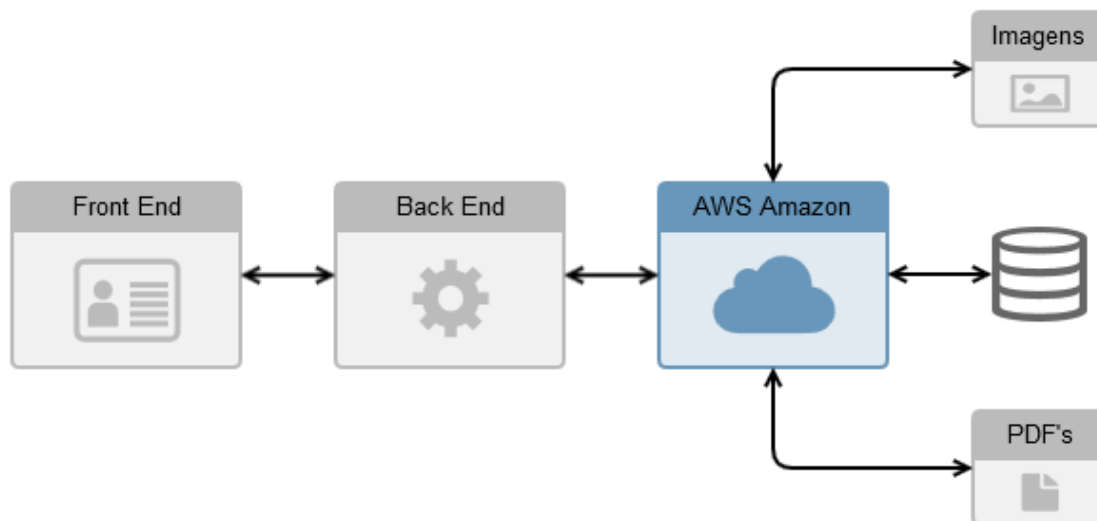


Ilustração 1 - Esquema geral do projeto

3.3.2 Diagramas de classes

A base de dados num projeto desta dimensão é essencial e requiere uma estrutura bem pensada. No nosso portal necessitamos de duas estruturas de base de dados diferentes. A estrutura principal que ao longo do tempo sofreu várias alterações, incluindo alterações por minha parte, na questão de performance de importação de produtos para o portal e garantia da melhor consistência entre as tabelas. Esta tabela (Ilustração 2) já existia com a maior parte de funcionalidades como multilingue, produtos, ficheiros, etc. Para guardar informação temporária (ver Ilustração 3) de produtos que estão a ser importados foi desenvolvida de raiz e garante o armazenamento desta informação e também informação sobre os erros de importação, que futuramente podem ser analisados para melhorar o serviço de importação para que haja menos

erros frequentes, visto que os registos de erros permanecem sempre na base de dados do sistema.

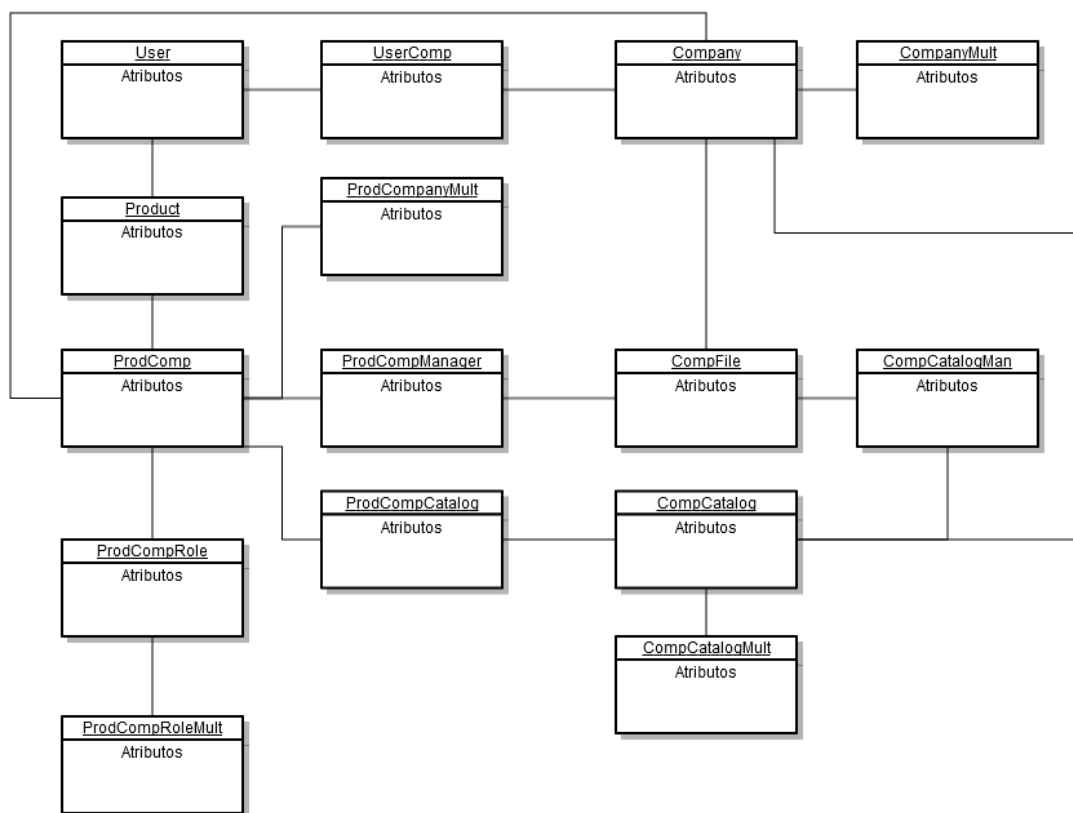


Ilustração 2 - Diagrama de Classes - Tabelas portal

A estrutura da Ilustração 2 corresponde a todas as necessidades do portal para que a informação seja lá guardada for facilmente apresentada ao utilizador. A estrutura garante vários pontos:

- Permite ao utilizador ter várias empresas;
- Uma empresa pode ser gerida por vários utilizadores;
- Um produto pode estar associado a várias empresas;
- Um produto suporta a tradução de campos para várias línguas;
- Um produto pode ter vários ficheiros (imagens e pdf's) associados;
- O mesmo produto em línguas diferentes pode ter ficheiros diferentes associados a ele;
- Um ficheiro pode ser usado em diferentes produtos;
- Um catálogo pode ter vários produtos e ficheiros associados a ele;
- Um catálogo suporta traduções para diferentes línguas;
- Linhas de catálogo diferentes podem ter os mesmos produtos ou ficheiros associados.

Com uma estrutura tão complexa conseguimos garantir todas as hipóteses necessárias para o negócio. Por ser uma estrutura complexa não quer dizer que seja complicado de a implementar, o mais complicado é conseguir chegar a uma solução que pode de forma simples englobar todos os problemas que possam surgir no futuro.

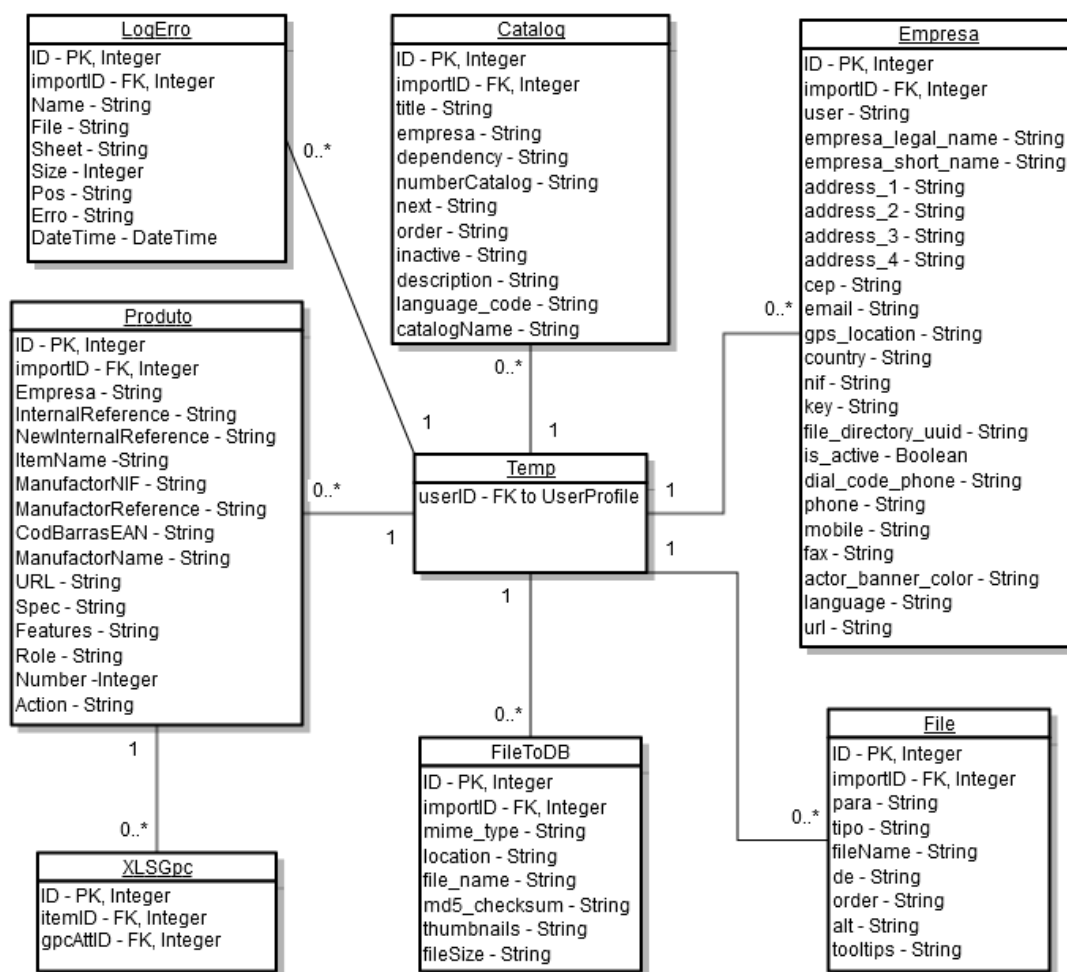


Ilustração 3 - Diagrama de Classes - Tabelas temporárias

3.3.3 Casos de uso

A aplicação tem três tipos de utilizadores (Ilustração 4). Um deles é o utilizador principal que gere os seus produtos, faz importação de ficheiros e ainda associa-os aos produtos, etc. O segundo tipo de utilizador são os que visualizam os produtos com ou sem os seus ficheiros (imagens, pdf's) associadas. E por fim o administrador do portal que consegue efectuar tudo o que faz o utilizador mais avançado e ainda enviar email para todos os utilizadores com as informação sobre os novos produtos que existem no portal. Por fim, o administrador pode também publicar os produtos no portal, isto é, fazer com que os produtos sejam visíveis para todos.

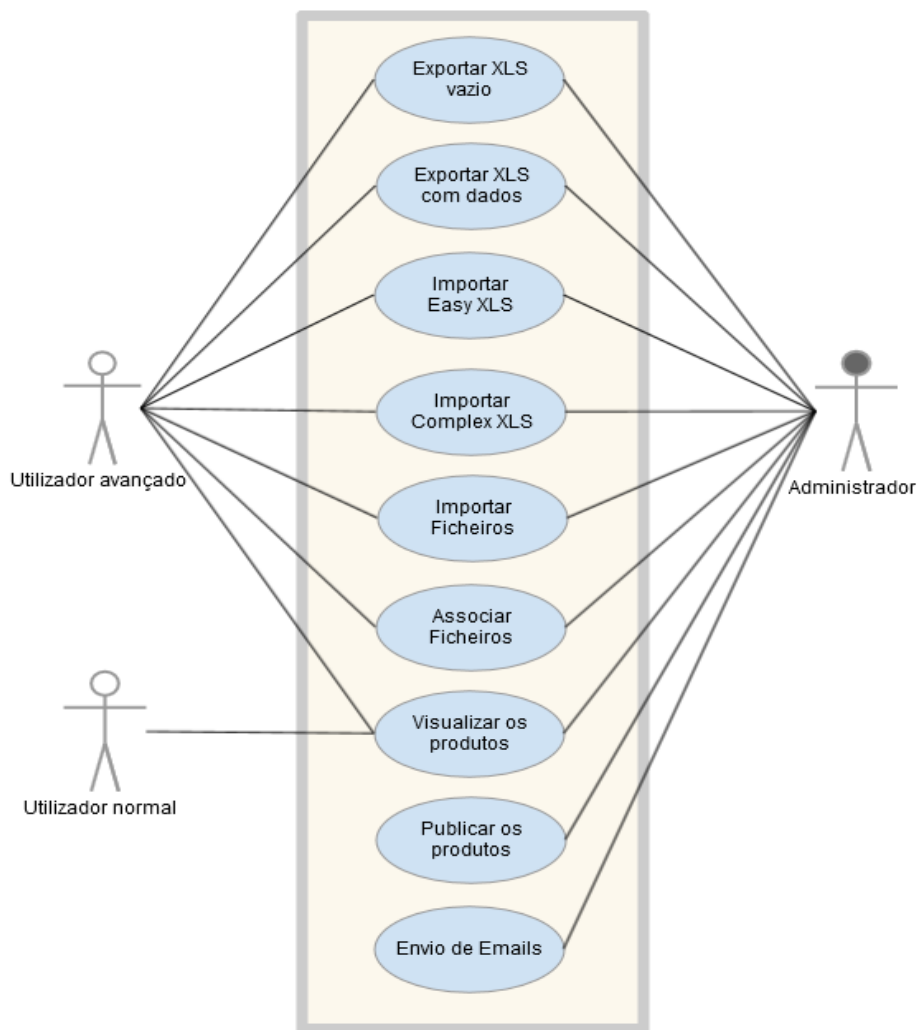


Ilustração 4 - Casos de uso - Gestão de produtos

3.4 Tecnologia usada

Ao escolher as tecnologias estava em conta o feedback delas e as ferramentas fornecidas para ter melhor performance e fácil integração. Durante o projeto o feedback dado sobre as tecnologias foi positivo. Estas tecnologias serão explicadas nos próximos pontos, assim como serão dados alguns exemplos de execução.

3.4.1 Django

Django [21] é uma ferramenta web opensource que utiliza a linguagem python [22] para realização de processos. A framework usa o modelo MVT (Model, View, Template) [23] que tem como a base principal o modelo MVC [24]. No model é definida a estrutura do projeto e é responsável pela execução de queries sobre uma base de dados. Na view é definida a informação que será apresentada ao utilizador, estruturado e enviado como um HTTP response. O Template por sua vez recebe a informação que já foi tratada e apresenta ao utilizador em formato programado, como por exemplo HTML/XML, etc.

A tecnologia Django é relativamente recente. Começou por lançar a sua primeira versão em 2005 e conta atualmente com uma versão estável 1.8. Para uma tecnologia tão recente ela já é muito utilizada por muitas empresas de pequena, média e grande dimensão, devido ao seu simples funcionamento e rapidez.

Criar um projeto em Django é simples [25], é necessário inicialmente instalar só o django e python e criar o seu próprio site:

- Começamos por construir o projeto executando na linha de comandos: `django-admin.py startproject helloworld`. Isto irá criar uma estrutura simples do projeto:
 - `manage.py` - o principal ficheiro para executar algo sobre a sua página;
 - `helloworld/settings.py` - contém todas as configurações da sua página, desde a configuração da base de dados até à declaração de variáveis do sistema;
 - `helloworld/urls.py` - é a construção lógica entre o código e o url da sua pagina.
- De seguida é necessário criar ficheiro `views.py` no seu projeto com o conteúdo do Código 1;
- Depois do conteúdo que irá ser apresentado ao utilizador é necessário fazer a ligação logica entre o nome do url e o código onde este esta situado. A ligação logica é exemplificada no Código 2.

```
from django.http import HttpResponseRedirect
def home(request):
    html = "<html><body>Hello world!</body></html>"
    return HttpResponseRedirect(html)
```

Código 1 - Criação da view. Hello world!

```
from django.conf.urls import patterns, url
urlpatterns = patterns('',
    url(r'^$', 'helloworld.views.home', name='home'),
)
```

Código 2 - Criação do URL. Hello world!

O presente exemplo não demonstra a criação de templates nem a configuração do site, visto que é um exemplo muito básico.

Para a execução do servidor e ver as suas aplicações a funcionar é necessário na linha de comandos executar o *python manager.py runserver* manager. Para entrar na sua aplicação num browser é necessário escrever o endereço e a porta onde o servidor está a correr, no nosso caso é *localhost:8000*, a porta por defeito do servidor Django é 8000. Ao arrancar o servidor, o django verifica o projeto todo e se houver algo de errado na estrutura ou mesmo no código é dado um feedback com o erro que levou o servidor a não conseguir arrancar. Com a correção do erro, o django verifica novamente o projeto e no caso de sucesso o servidor é arrancado para utilizar a sua aplicação.

3.4.2 Python

Python é uma linguagem de programação muito antiga e muito usada hoje em dia. É simples, prática e rápida em execução das suas tarefas. Para o correto funcionamento da linguagem é necessário garantir a tabulação de 4 espaços ao estar a entrar dentro de classes, métodos, if's, for's, etc. Python consegue receber vários valores nos seus métodos e o mesmo acontece ao retornar os valores de um determinado método. As variáveis não precisam de ser declaradas, mas sim usadas diretamente. No entanto, cada variável depois de ser atribuída com um valor tem o seu tipo. Alguns dos exemplos pode ver no Código 3.

```
class newClass(object):
    def metodo(self, paramReg, param2):
        var = 0
        for reg in paramReg:
            if(reg == param2):
                var += 1
            else:
                print "Hello world " + var
        return var, paramReg
```

Código 3 - Exemplos python

Python contem muitas bibliotecas que estão constantemente a serem atualizadas para garantir aos utilizadores as últimas soluções e uma execução rápida. É possível fazer de tudo em python e é muito simples de utilizar. Basta instalar a biblioteca pretendida e fazer importação da mesma para o projeto onde é necessário utilizá-la.

Para o debugging no django utilizamos biblioteca pdb do python, que tem uma utilização bastante simples. No código onde é necessário fazer o debug basta escrever o seguinte Código 4 e na linha

de comandos onde estamos a executar o nosso servidor aparece o cursor. Durante o debug é possível visualizar todas as variáveis que foram executadas até ao momento do debug e alterá-las para ver como irá ser executado o resto do código.

```
import pdb; pdb.set_trace()
```

Código 4 - Python debugging

A simplicidade e flexibilidade desta linguagem é que nos levou à escolha da mesma. Através desta linguagem conseguimos executar processos como tratamento de imagens, leitura de ficheiro xls, etc, com uma boa performance e com pouco tempo de implementação.

3.4.3 Postgresql

PostgreSQL [26] é um sistema de utilização de base de dados relacionais, opensource, segue standards SQL e é 100% ACID (Atomicity, Consistency, Isolation, Durability) [27] . PostgreSQL é suportado por vários sistemas operativos. De entre os mais utilizados estão Windows, Linux, Mac OS, etc. Postgres tem grandes possibilidades de guardar a informação, como por exemplo, base de dados infinita, uma tabela com um máximo de 32 Tb, etc. [28]

Postgresql é facilmente integrada nos projetos Django e tem uma construção de base de dados muito prática, utilizando os modelos de Django para construir a sua estrutura. Depois da estrutura estar criada é necessário fazer a sincronização da estrutura criada com a estrutura já existe. Para isso é necessário configurar o Django para receber a base de dados Postgresql com um user previamente criado. Assim, depois disso é executado *python manage.py syncdb* que faz a ligação e atualiza todas as tabelas da base de dados. O acesso aos registos da base de dados é muito simples, como é exemplificado no Código 5. Os valores que a query devolve vêm numa lista de registos e em python tem um acesso bastante simples.

```
#Selecionar todos os registos que tenham como nome Anton e que não sejam  
do staff, ordenadopor id decrescente. select_related - Seleção de  
tabelas em cascata.  
Exemplo1.objects.select_related().filter(nome = "Anton", staff = False)  
.order_by("-id")  
  
#Seleção de todos os registos que tenha nome Anton na tabela Exemplo2 e  
em cascata a tabela Exemplo3  
Exemplo2.objects.select_related("Exemplo3").filter(nome = "Anton")  
  
#Seleção de todos os registos que na tabela exemploB tenha nome Anton  
ExemploA.objects.select_related().filter(exemploB__nome = "Anton")
```

Código 5 - Exemplo - Execução queries

A base de dados Postgresql foi escolhida porque é uma bases de dados gratuita e de bom funcionamento com grande volume de dados e com boa documentação. É uma das partes essenciais para um portal que irá ter muitos clientes, nacionais e internacionais, com grande gama de produtos.

4 Implementação

4.1 Exportação

Pode existir duas formas de exportação de diferentes ficheiros XLS (Simple XLS e Complex XLS):

- A primeira forma é a exportação do ficheiro simples com apenas informação sobre o título das colunas. Estes ficheiros são necessários para que o utilizador consiga preencher corretamente o ficheiro. Na Ilustração 5 está explicado o processo de exportação do ficheiro xls.
- A segunda forma é um ficheiro igual ao ficheiro simples só que já com informação preenchida, de todos os produtos do utilizador. Este tipo de exportação é necessário para utilizador poder facilmente extrair os seus produtos para ficheiro xls caso precise de utiliza-los futuramente.

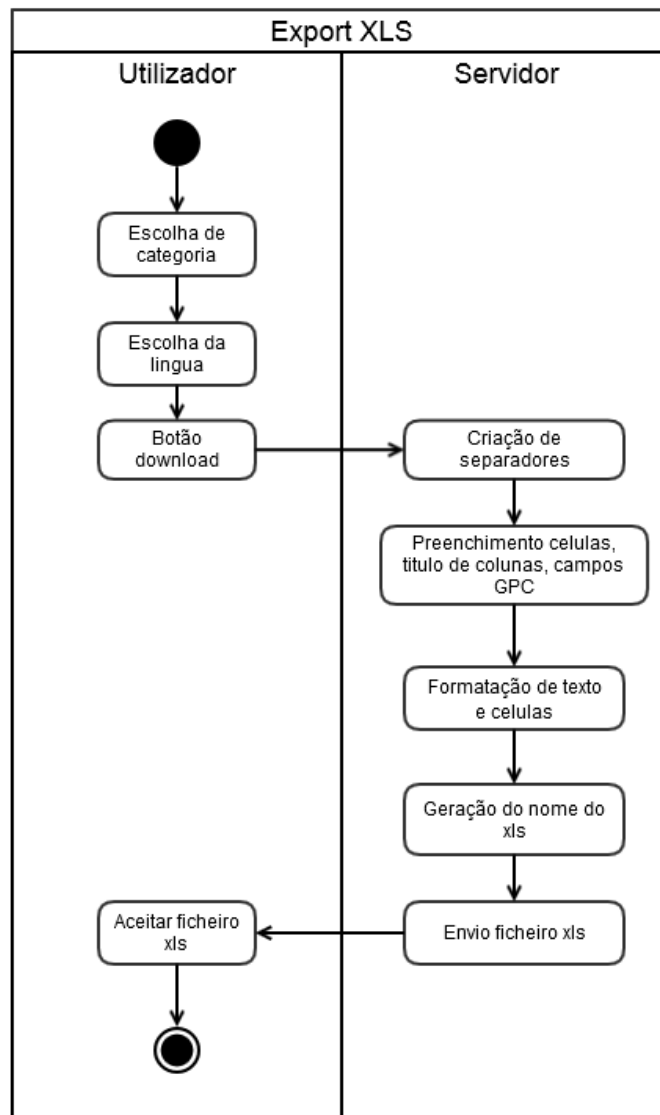


Ilustração 5 - Diagrama de atividade - Exportação XLS

4.1.1 Simple XLS (simples)

Para exportar o ficheiro Simple XLS é necessário escolher a categoria GPC [29] em qual vão ser inseridos os produtos e a língua em qual pretende exportar os seus produtos (Ilustração 6). Depois de escolher a categoria correta e a língua correta basta carregar no botão para que o ficheiro Simple XLS seja transferido para o computador do utilizador.

Seleccione a categoria para os seus produtos*

Audiovisual e Fotografia v Fotografia e Ótica v Ótica v Microscópios v

A categoria que escolheu: **Audiovisual e Fotografia** > **Fotografia e Ótica** > **Ótica** > **Microscópios** ✓

Configure o carregamento dos seus produtos

Selecione a língua para os produtos que vai carregar

Português v

Download de XLS

Ilustração 6 - Export XLS - Fase inicial

O ficheiro criado segue algumas normas para que seja interpretado de uma melhor forma por utilizador. E também esconde algumas partes relevantes para a aplicação e desnecessários para o utilizador.

O primeiro separador chama-se 'ITEM', contém 8 colunas estáticas e colunas dinâmicas que servem para guardar informações sobre os produtos do utilizador.

Na primeira linha, coluna A é guardado o ID do brick (ver 4.2.11 GPC categorias) das categorias GPC para que seja identificado a que categoria vão ser associados os produtos. Na coluna B da mesma linha é guardada a língua em que os produtos serão importados, no formato de duas letras, como por exemplo 'PT'. Essa linha é escondida do utilizador para não ser alterada por engano, pois esta informação só é relevante ao sistema de importação.

A segunda linha contém a informação da categoria escolhida por utilizador. Essa linha é simplesmente informativa para que o utilizador não tenha dúvidas da categoria que esta a preencher.

A terceira linha também é só para utilização interna que guarda informação dos atributos de cada brick, isto serve para o sistema de importação saber qual é o atributo a ser lido, visto que todos os atributos do GPC são diferentes.

Na quarta linha estão os nomes de cada coluna que identificam a informação a preencher de cada produto. Os nomes são divididos por três grupos diferentes:

1. Obrigatórios. O fundo desses campos é um laranja claro e contem uma estrela no final do nome do campo, para que for mais familiar ao utilizador identificar os campos obrigatórios. Estes campos correspondem as necessidades mínimas do portal e são os seguintes:
 - a. Nome do produto*. O nome do produto para transmitir a informação simples e útil;
 - b. Características*. Descrição simples e breve do produto a importar;

- c. Tag do produto*. Uma tag para que seja facilmente encontrado na pesquisa do portal;
 - d. Referência interna do produto. A referência do produto tem de ser campo único no xls e é único em todos os produtos do cliente, é o campo que diferencia todos os produtos entre si.
2. Complementares. O fundo desses campos é um laranja mais escuro do que nos obrigatórios. Os campos são os seguintes:
 - a. Código de barras. É guardado como o segundo campo único do produto, caso este campo não seja preenchido é gerado automaticamente um código de 16 bits para identificação do produto. Ao introduzir o código este será verificado pela uma biblioteca python que verifica a validação do código de barras e devolve um resultado que será tratado da forma correta;
 - b. Marca. Marca do produto a ser importado;
 - c. Modelo. Modelo do produto a ser importado;
 - d. Especificações. Onde podemos indicar partes mais técnicas do produto
3. Atributos GPC. São campos dinâmicos que são gerados automaticamente por categoria GPC. Estes campos são gerados conforme o tipo de categoria que está a ser exportado (brick). No sistema é identificado o número de atributos que contém o determinado brick e esse número de atributos com os seus nomes serão acrescentados nas colunas do XLS. Todos os atributos contém os seus valores, que serão geridos da 5ª linha para baixo e são disponibilizados numa dropdown.

The screenshot shows an Excel spreadsheet with columns A through L. Row 1 contains '10002193' in column A and 'pt' in column B. Row 2 contains a long category string: 'Category: Móveis e Artigos Mobiliários para Casa e Escritório >> Mobiliário para Casa e Escritório >> Assentos para Casa e Escritório >> Cadeiras e Bancos para Casa e Escritório (Não Elétricas)'. Row 3 contains '20001708' in column F and '20001709' in column G. Row 4 is the header for the data section, with columns: 'Nome do produto*', 'Características*', 'Tag do produto*', 'Código de barras', 'Referência interna (Rodas)', 'Utilização ou Aplicação Alvo', 'Marca', 'Modelo', and 'Especificações'. Row 5 is the first data row, where the 'Utilização ou Aplicação Alvo' column has a dropdown menu open, showing a list of options: 'CASA DE BANHO', 'QUARTO', 'CANTINA', 'SALA DE JANTAR', 'ENTRADA', 'COZINHA', and 'SALÃO'. The 'CANTINA' option is currently selected. The spreadsheet interface includes a menu bar with 'ITEM', 'FILE TO', and 'AJUDA', and a status bar at the bottom.

Ilustração 7 - Ficheiro Simple XLS. Separador ITEM

O segundo separador contém informação necessária para poder associar ficheiros (imagens e pdf's) aos produtos que estão inseridos no primeiro separador. Existem 6 colunas que são necessárias para processar e completar os ficheiros associados, entre quais 4 delas são obrigatórias para poder conseguir associar os ficheiros certos aos produtos certos e chama-se 'FILE TO':

1. Tipo, é um campo do tipo lista que contém informação sobre o tipo do ficheiro que vai ser importado, imagem (IMG) ou pdf (PDF). Está disponível ao utilizador através duma dropdown.
2. Nome do ficheiro, é o ficheiro que queremos associar ao produto. No nome do ficheiro a extensão não é identificada, pois no campo descrito acima já identificamos o ficheiro que está a ser transferido. Mas também esta decisão foi feita a pensar no utilizador final que por sua vez nem sempre consegue identificar que tipo de extensão tem determinado ficheiro.

3. Referência interna do produto identifica a que produto pretende-se associar o ficheiro.
4. Ordem, serve para identificar pela qual ordem serão apresentados as imagens ou pdf's no portal, pois por vezes existe a necessidade de mudar a imagem principal e será feito de uma forma mais rápida trocando só o numero da ordem das imagens.

Os outros dois campos são opcionais e permitem especificar melhor o ficheiro que está a ser transferido:

1. Alt, é um tipo de campo que apresenta uma alternativa ao nome de um ficheiro (é mais usado para os ficheiros pdf).
2. Tooltip, é um campo que será mais usado para imagens que contêm a descrição muito simples e rápida de uma imagem e/ou produto. Ao ter o cursor do rato por cima da imagem esse campo ficará visível.

	A	B	C	D	E	F
1	Tipo (IMG/PDF)*	Nome do ficheiro*	Referência interna do produto*	Ordem*	Alt	Tooltip
2	<input type="text"/>					
3	IMG					
4	PDF					
5						
6						
7						
8						
9						

Ilustração 8 - Ficheiro Simple XLS. Separador FILE TO

O terceiro separador contém informações básicas para preenchimento do ficheiro XLS, para que o utilizador possa, sempre que é preciso, consultar a maneira correta de inserção de produtos.

Todos os nomes de colunas e as suas descrições vêm na língua de perfil do utilizador, caso o utilizador tenha uma língua para qual não exista tradução então a língua por omissão é inglês.

No final da geração do ficheiro completo, é atribuído um nome ao ficheiro que contém o nome da empresa, data e o ID da categoria GPC e tem o seguinte formato: EmpX140503_10219.xlsx.

A geração do ficheiro é feita através duma biblioteca python que perante o estudo deu melhores resultados em relação a flexibilidade de criação do formato do ficheiro XLS.

A biblioteca é `xlsxwriter` [30], permite criar os ficheiros `xlsx` e fazer vários tipos de formatações:

- Escolha do tipo de letra;
- Definir o tamanho da letra;
- Propriedades da letra (Negrito, Itálico, Sublinhado, etc);
- Definir a cor de fundo utilizando as cores RGB;
- Definir a largura de colunas e linhas;
- Definir valores em lista, apresentados numa dropdown;
- Criar separadores e definir os nomes;

- Atribuir o nome do ficheiro;
- Criar fórmulas;
- Etc.

Esta biblioteca foi bastante útil para que o ficheiro gerido ficasse mais adaptado ao utilizador.

4.1.2 Complex XLS(simples)

O ficheiro XLS mais complexo foi criado para conseguir importar mais informação do utilizador e dos seus produtos, aquilo que não é possível fazer com Simple XLS.

Com o Complex XLS é possível importar uma empresa, estrutura de catálogo, produtos, ficheiros associados aos produtos, fornecedores de cada produto e alguns parâmetros. O Complex XLS permite preencher todos os detalhes em relação aos produtos com todo o detalhe.

O Complex XLS contém 7 separadores que agrupa dentro de si toda a informação necessária. Depois de importar o xls corretamente preenchido, o catálogo do utilizador fica completamente pronto para disponibilizar online.

No primeiro separador será preenchida a informação sobre empresa de um utilizador. Existe 9 campos obrigatórios para o funcionamento mínimo de uma empresa no portal, como por exemplo: nome do utilizador a qual vai pertencer a empresa, morada da empresa, nome da empresa, etc. Também existe campos opcionais para completar a informação da empresa, como por exemplo: telefone da empresa, url da empresa, morada mais detalhada, etc.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	User Name	Cor banner	Nome legal	Short Name	País	NIF/VAT	Morada linha 1	Morada linha 2	Morada linha 3	Morada linha 4	Morada CEP	Localização GPS	Cod indicador telefonico	Telemóvel	Telefone	Fax	E-mail	URL
2																		
3																		
4																		
5																		
6																		

Ilustração 9 - Ficheiro Complex XLS. Separador ACTOR

No segundo separador pretende-se agrupar toda a informação relativamente ao catálogo de uma empresa. Este catálogo pode ser dividido por vários níveis, sendo o máximo de 6 níveis, para armazenar a informação da empresa de uma forma mais precisa. Por cada linha do catálogo é possível dar o nome da linha do catálogo, título e a informação sobre o mesmo. Para distinguir qual é a estrutura do catálogo, vão ser utilizados os campos de número de catálogo e a dependência, no qual é necessário indicar o número do catálogo e do qual o mesmo depende.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Número do catálogo	Dependência do catálogo	Nome do catálogo	Título do catálogo	Informação do catálogo	
2												

Ilustração 10 - Ficheiro Complex XLS. Separador CATALOG TREE

O terceiro separado contém a informação sobre os produtos da empresa. Neste separador podemos inserir, modificar ou apagar os produtos. Para isso utilizamos uma coluna onde

indicamos o tipo da operação que se pretende fazer. Esta informação esta numa dropdown (IGNORE, ADD, MODIFY, DELETE) em cada célula da coluna "Ação". O resto das colunas são geradas normalmente como células simples. O resto dos campos tem informação sobre o produto, como por exemplo: nome, descrição, especificações, código de barras, dados do fornecedor, etc.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Ação	UPC do ITEM (se existir)	Referência interna actual ITEM	Nova referência interna ITEM	Nome ITEM	Características do ITEM	Especificações do ITEM	Papel da empresa nest ITEM (ACTOR ROLE)	Nº do CATALOG	Referência do fabricante para este ITEM	NIF/VAT fabricante deste ITEM	Nome do fabricante deste ITEM	URL do fabricante deste ITEM
2													
3	IGNORE												
4	ADD												
5	MODIFY												
6	DELETE												
7													

Ilustração 11 - Ficheiro Complex XLS. Separador ITEM

O quarto separador contem a informação em relação a ficheiros que vão ser associados a produtos, empresa e catálogo. Neste separador existe duas colunas com células especiais, que são:

- Destino do ficheiro, onde indicamos a que vai pertencer o ficheiro. O ficheiro pode ser associado aos produtos, catálogos e a empresa (logotipo).
- Tipo de ficheiro que está a ser associado, que pode ser do tipo IMG, ou seja, imagem ou do tipo PDF.

Ambos os campos descritos acima são células dropdown com valores predefinidos.

O resto das células são normais e servem para indicar o nome do ficheiro, a referência do produto, número do catálogo ou o nome da empresa a que pretende associar determinado ficheiro. Também tem a ordem para casos em que existe mais do que um ficheiro por produto e os habituais Alt e Tooltip.

	A	B	C	D	E	F	G
1	Para (ACTOR, CATALOG ou ITEM)	Tipo de ficheiro - Imagem ou PDF	Nome do Ficheiro	Destino (ref ITEM, nº CATALOG, shortname do ACTOR)	Nº de ordem dentro do mesmo destino	Alt / Título	Tooltip
2							
3							
4	ITEM						
5	CATALOG						
6	ACTOR						

Ilustração 12 - Ficheiro Complex XLS. Separador FILE TO

No quinto separador é guardada a informação do fornecedor de cada produto. Todas as células deste separador são simples e todas elas são opcionais.

	A	B	C	D
1	Referência interna ITEM	Referência do fornecedor	NIF/VAT do fornecedor	Nome do fornecedor
2				
3				
4				
5				

ACTOR CATALOG_TREE ITEM FILE TO **SUPPLIERS** PARAMETERS AJUDA

Ilustração 13 - Ficheiro Complex XLS. Separador SUPPLIERS

O último separador contém informação necessária para futura importação. A língua em qual o catálogo vai ser importado e o ID da empresa, este campo só serve para casos de alterações dos produtos.

	A	B	C	D	E	F	G
1	Idioma do Catálogo	ACTOR ID					
2							

ACTOR CATALOG_TREE ITEM FILE TO SUPPLIERS **PARAMETERS** AJUDA

Ilustração 14 - Ficheiro Complex XLS. Separador PARAMETERS

O ficheiro Complex XLS é exportado na língua do perfil do utilizador. Caso o utilizador tenha a língua do perfil para qual não existe a tradução então vem na língua por omissão, inglês.

A biblioteca usada para a exportação do ficheiro Complex XLS é a mesma que utilizamos para o Simple XLS, `xlsxwriter`.

A categoria GPC em que vão ser guardados os produtos é uma categoria default, porque como a quantidade de informação e produtos é elevada e na maioria dos casos os produtos são para diferentes linhas de catálogo, decidiu-se não especificar o tipo de catálogo para este caso de XLS.

4.1.3 Simple XLS (com dados)

Ao exportar o ficheiro Simple XLS com dados obtemos os produtos e toda a informação sobre eles de uma determinada empresa.

Para a exportação com dados é necessário escolher a língua de produtos que pretendemos exportar. A exportação vem automaticamente com a categoria GPC default, porque muitos produtos normalmente têm mais do que uma categoria. As células dropdown no separador FILE TO são na mesma do tipo lista com valores predefinidos, mas depois da exportação com dados esses valores ficam preenchidos.

O nome do ficheiro exportado contém o nome da empresa e a data da exportação, exemplo: `AntonzaCo_070520150114.xlsx`.

4.1.4 Complex XLS (com dados)

O ficheiro Complex XLS contém toda a informação da empresa, estrutura do catálogo, produtos, ficheiros, fornecedores e parâmetros.

Para a exportação é necessário escolher a língua dos produtos que pretendemos exportar. Ao exportar os separadores ficam de seguinte forma:

- ACTOR é preenchido por toda a informação que já esteja guardada no portal;
- CATALOG TREE é preenchido com toda a lista de linhas de catálogo e as suas descrições. Para gerar a árvore das linhas de categorias foi usada uma estrutura de dicionário de python para guardar todas as linhas com as suas dependências. Isto facilita a inserção da informação da árvore do catálogo no xls;
- ITEM é preenchido na totalidade com os seus dados. São preenchidos todos os campos menos os de código de barras, caso este seja um stamp gerado automaticamente. A coluna de Ação vem preenchida com valor "IGNORE" para que estes campos não serem lidos na próxima importação. Se o utilizador pretender que os produtos sejam lidos então na dropdown tem de escolher uma das outras opções;
- FILE TO exporta a informação de todos os ficheiros que existem no portal e que estejam associados aos produtos que estão a ser exportados. Na coluna de destinatário, as células continua a ser do tipo lista com valores predefinidos. As células de tipo de ficheiro também são do tipo lista (IMG e PDF) com valores já preenchidos, conforme o ficheiro for;
- SUPPLIERS é preenchido com a informação dos fornecedores;
- PARAMETERS é preenchido com a língua na qual estamos a exportar os produtos e o ID da empresa para necessidades futuras.

O nome do ficheiro exportado contém o nome da empresa e a data da exportação, exemplo: AntonzaCo_070520150114.xlsx.

4.2 Importação

As técnicas de importação são uma parte essencial do projeto. A importação garante que os produtos do utilizador sejam rapidamente transferidos para o portal de forma correta. Além da importação de produtos do utilizador, também explico outros tipos de processos de importação que foram desenvolvidos durante o projeto.

A importação dos produtos para o portal é feita em duas partes(ver Ilustração 15). Na primeira parte, o utilizador verifica todos os produtos inseridos no xls com ou sem ficheiros, confirma se as associações que foram feitas estão como foi previamente planeado e também é apresentado um relatório de erros, caso estes existam. A segunda parte da importação serve para confirmar o que foi inserido pelo utilizador e guardar os produtos no portal.

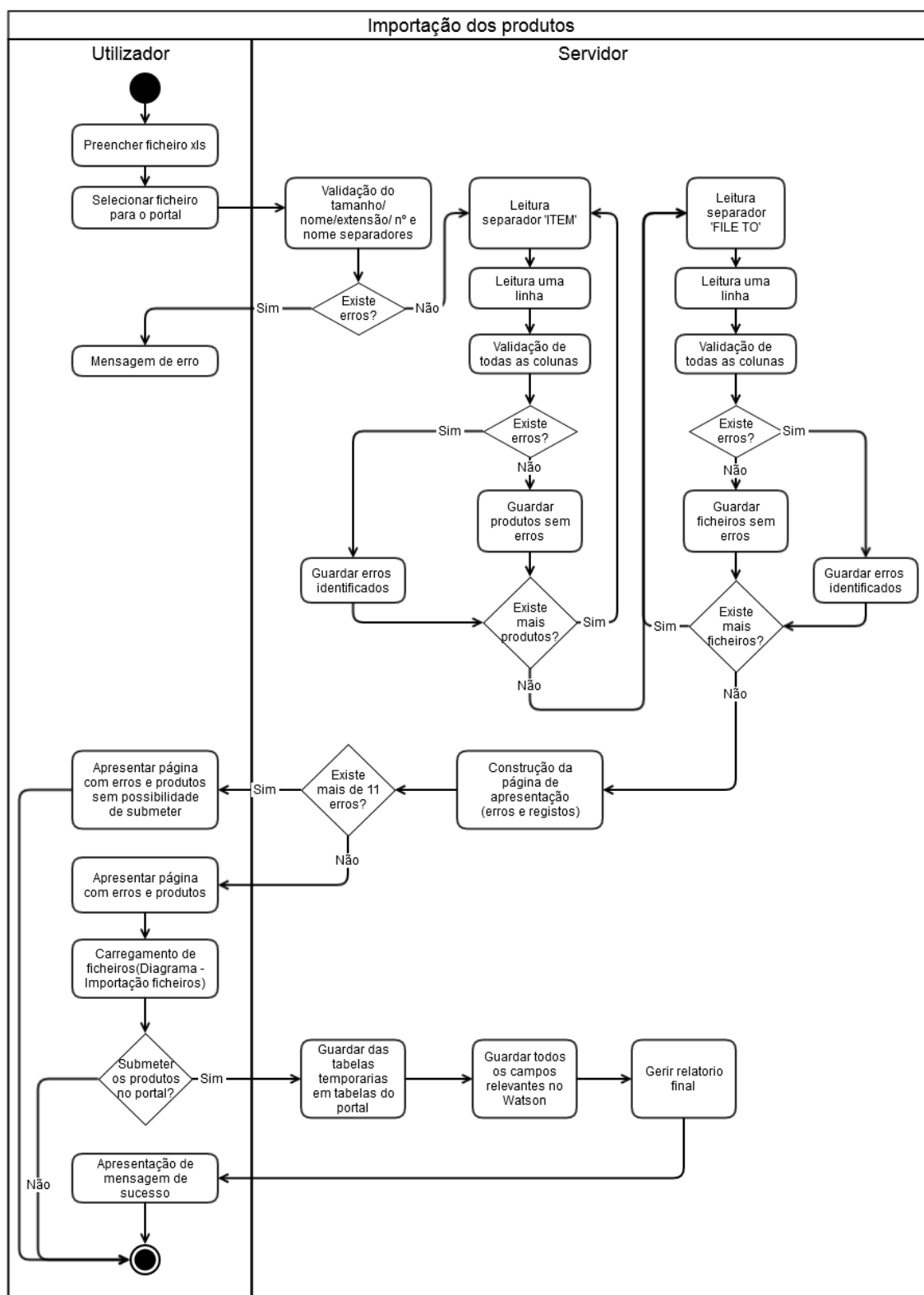


Ilustração 15 - Diagrama de Actividade - Importação Simple XLS

4.2.1 Importação Simple XLS - Carregamento temporário

A primeira parte deste processo começa pelo utilizador escolher o ficheiro xls que criou com seguimento de todas as regras indicadas no ficheiro. Este passo é importante para que o ficheiro não apresente erros.

Para escolher o ficheiro utilizei a biblioteca do django *form* [31] que facilmente guarda o ficheiro em memória para que possa ser facilmente processado.

Para a leitura do ficheiro XLS é utilizada a biblioteca python *xlrd* [32]. É uma forma simples e prática de abrir o ficheiro XLS e ler toda a informação presente no mesmo. A biblioteca deteta o número de separadores que existem no ficheiro e lê todas as linhas e colunas. Por outro lado, apenas é lido até ao limite de linha/coluna máxima do separador. As células lidas pela biblioteca são interpretadas em dois tipos de dados, String ou Double, por isso ao ler as células é necessário converter alguns valores para Inteiros para guardar corretamente na base de dados.

A verificação do ficheiro começa por simples verificações sem estar a aplicar a abertura do ficheiro em si e caso existam erros é logo descartado na fase inicial de todo o processo. As primeiras verificações são feitas à extensão do ficheiro, a qual deve ser obrigatoriamente do tipo xls ou.xlsx. Por outro lado, é verificado se o ficheiro é superior a 0 Kb e inferior a 30 Mb.

Depois de abrir o ficheiro em memória, começamos por executar as confirmações mais detalhadas:

- Número de separadores e os seus nomes;
- Número das colunas por cada separador, sendo que o número de colunas do separador ITEM pode ser diferente, conforme os atributos do GPC;
- Língua do em que os produtos vão ser inseridos (esta célula está escondida do utilizador);
- ID do GPC;
- ID's dos atributos da categoria GPC, caso existam.

Caso estes pontos correspondam às necessidades, então irá começar a validação de todas as linhas de cada separador.

Inicialmente é lido o primeiro separador ITEM, como mostra a Ilustração 16. Posteriormente são lidas todas as linhas, coluna a coluna e validados os campos que foram inseridos pelo utilizador. Caso algum dos campos não corresponder às validações então é adicionado um erro numa tabela de erros que vão ser apresentados ao utilizador. Por fim, o utilizador pode corrigir o ficheiro e importar mais uma vez. Se preferir, pode transferir para o portal apenas os produtos que não geraram nenhum tipo de erro.

Se pelo menos um dos valores não corresponder às validações, este fica descartado e na corrente importação não será possível a sua importação para o portal.

	A	B	C	D	E	F	G	H	I	J
2	Category: Móveis e Artigos Mobiliários para Casa e Escritório >> Mobiliário para Casa e Escritório >> Assentos para Casa e Escritório >> Cadeiras e Bancos para Casa e Escritório (Não Elétricas)									
4	Nome do produto*	Características*	Tag do produto*	Código de barras	Referência inte	Rodas	Utilização ou Aplicação Alvo	Marca	Modelo	Especificações
5	Produto CA-1	Características CA-1	Produto teste		CA - 0001	NÃO	CASA DE BANHO	Marca teste	Modelo teste	Especificações CA1
6	Produto CA-2	Características CA-2	Produto teste		CA - 0002	NÃO	QUARTO	Marca teste	Modelo teste	Especificações CA2
7	Produto CA-3	Características CA-3	Produto teste		CA - 0003	SIM	CANTINA	Marca teste	Modelo teste	Especificações CA3
8	Produto CA-4	Características CA-4	Produto teste		CA - 0004	NÃO IDENTIFICADO	SALA DE JANTAR	Marca teste	Modelo teste	Especificações CA4
9	Produto CA-5	Características CA-5	Produto teste		CA - 0005	NÃO	SALÃO	Marca teste	Modelo teste	Especificações CA5
10	Produto CA-6	Características CA-6	Produto teste		CA - 0006	SIM	UNIVERSAL	Marca teste	Modelo teste	Especificações CA6
11	Produto CA-7	Características CA-7	Produto teste		CA - 0007	NÃO IDENTIFICADO	CANTINA	Marca teste	Modelo teste	Especificações CA7
12	Produto CA-8	Características CA-8	Produto teste		CA - 0008	NÃO IDENTIFICADO	NÃO CLASSIFICADO	Marca teste	Modelo teste	Especificações CA8
13	Produto CA-9	Características CA-9	Produto teste		CA - 0009	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA9
14	Produto CA-10	Características CA-10	Produto teste		CA - 0010	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA10
15	Produto CA-11	Características CA-11	Produto teste		CA - 0011	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA11
16	Produto CA-12	Características CA-12	Produto teste		CA - 0012	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA12
17	Produto CA-13	Características CA-13	Produto teste		CA - 0013	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA13
18	Produto CA-14	Características CA-14	Produto teste		CA - 0014	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA14
19	Produto CA-15	Características CA-15	Produto teste		CA - 0015	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA15
20	Produto CA-16	Características CA-16	Produto teste		CA - 0016	NÃO IDENTIFICADO	UNIVERSAL	Marca teste	Modelo teste	Especificações CA16

Ilustração 16 - Import Simple XLS - Separador ITEM - Preenchido

Os produtos que não apresentarem nenhum tipo de erro serão guardados numa tabela temporária para que sejam importados para o portal, depois do utilizador confirmar os produtos. A tabela temporária, além dos campos que estão presentes no ficheiro xls, contém o ID do utilizador que está a fazer a transferência. Este ID existe para diferenciar os produtos na tabela temporária no caso de existir mais do que um utilizador a importar os produtos. Alguns dos campos serão guardados em tabelas diferentes. Estes campos dinâmicos, são guardados em tabelas diferentes uma vez que tem uma estrutura diferente e não seria prático se fossem inseridos na mesma tabela dos itens.

De seguida é analisado o segundo separador (FILE TO), como mostra a Ilustração 16. Este separador segue a mesma teoria que o separador ITEM e caso cada linha passar as validações será guardada numa tabela temporária para posteriormente ser transferida para as tabelas do portal. Estes dados, presentes na tabela temporária, também servirão para o upload dos ficheiros que vão ser associados aos produtos.

	A	B	C	D	E	F
1	Tipo (IMG/PDF)*	Nome do ficheiro*	Referência interna	Ordem*	Alt	Tooltip
2	IMG	CA - 0001	CA - 0001	1	Alt Produto CA-1	Tooltip Produto CA-1
3	IMG	CA - 0002	CA - 0002	1	Alt Produto CA-2	Tooltip Produto CA-2
4	IMG	CA - 0003	CA - 0003	1	Alt Produto CA-3	Tooltip Produto CA-3
5	IMG	CA - 0004_01	CA - 0004	1	Alt Produto CA-4	Tooltip Produto CA-4
6	IMG	CA - 0004_02	CA - 0004	2	Alt Produto CA-5	Tooltip Produto CA-5
7	IMG	CA - 0005	CA - 0005	1	Alt Produto CA-6	Tooltip Produto CA-6
8	IMG	CA - 0007	CA - 0006	1	Alt Produto CA-7	Tooltip Produto CA-7
9	IMG	CA - 0008	CA - 0007	1	Alt Produto CA-8	Tooltip Produto CA-8
10	IMG	CA - 0009	CA - 0008	1	Alt Produto CA-9	Tooltip Produto CA-9
11	IMG	CA - 0010_01	CA - 0009	1	Alt Produto CA-10	Tooltip Produto CA-10
12	IMG	CA - 0010_02	CA - 0009	2	Alt Produto CA-11	Tooltip Produto CA-11
13	IMG	CA - 0011	CA - 0010	1	Alt Produto CA-12	Tooltip Produto CA-12
14	IMG	CA - 0012	CA - 0011	1	Alt Produto CA-13	Tooltip Produto CA-13
15	PDF	CA - 0013	CA - 0012	1	Alt Produto CA-14	Tooltip Produto CA-14
16	PDF	CA - 0014	CA - 0013	1	Alt Produto CA-15	Tooltip Produto CA-15
17	PDF	CA - 0015	CA - 0014	1	Alt Produto CA-16	Tooltip Produto CA-16
18	PDF	CA - 0016	CA - 0015	1	Alt Produto CA-17	Tooltip Produto CA-17
19	PDF	CA - 0017	CA - 0016	1	Alt Produto CA-18	Tooltip Produto CA-18

Ilustração 17 - Import Simple XLS - Separador FILE TO - Preenchido

Quando todas as linhas dos dois separadores forem lidas com sucesso, será gerado o relatório da importação no qual é possível verificar todos os produtos que foram inseridos no xls, assim como todos os nomes dos ficheiros e que tipos de ficheiros serão adicionados aos produtos. Além disso é apresentado em que língua vão ser guardados os produtos no portal, que tamanho tem o ficheiro xls, o nome do ficheiro xls e também a lista de erros com descrição detalhada de cada um, como mostra a Ilustração 18.

Relatório de upload

Nome do ficheiro

ImportXLS_10000785.xlsx

Tamanho do ficheiro

7638 kb

Língua do ficheiro:

pt

Estado do XLS

O ficheiro XLS está correto.

Nº de erros:

0

Seleccionar e fazer upload de ficheiros

Browse...

No files selected.

Tipos de ficheiro permitidos: PDF, JPG e JPEG

Adicionar ficheiros

Ver detalhes do relatório

Ilustração 18 - Importação Simple XLS - Relatório

No caso de existir ficheiros para associar aos produtos é necessário que estes sejam transferidos (4.2.8 Importação ficheiros) para o portal, onde serão validados e, no caso de sucesso, serem associados aos produtos.






Depois de carregar o ficheiro xls e todos os ficheiros que se pretende associar aos produtos, é possível ainda, na fase inicial, juntar os produtos com os ficheiros que foram transferidos, visualizando assim o produto e a imagem associada como mostra a Ilustração 19.

N	Product name	Brief description	Keyword	Bar code	Internal reference	Brand	Model	Additional description
1	Produto 1	Características do produto	Tag produto 1	6a441805-71da-474b-af09-006e06815d9e	lprod51	Marca do produto	Modelo do produto	Especificações do produto

GPC Attributes

Internal Reference(s) do Ficheiro para Attributs
lprod51 NÃO CLASSIFICADO

Images

N	Type	File Name	Item	Order	File	Thumbnail	Preview
1	IMG	lprod51	lprod51	1	lprod51	lprod51	
2	IMG	lprod52	lprod51	2	lprod52	lprod52	
3	IMG	lprod53	lprod51	3	lprod53	lprod53	
4	PDF	lprod54	lprod51	1	lprod54	lprod54	
5	PDF	lprod55	lprod51	2	lprod55	lprod55	

Verify if the file is correct and press Submit.

Ilustração 19 - Importação XLS - Apresentação de produtos

Caso esteja tudo correto existe dois procedimentos diferentes:

1. Cancelar o processo e desta forma todas as tabelas temporárias e diretorias em que estão guardados os ficheiros serão apagadas;
2. Seguir com importação dos produtos com os respetivos ficheiros para o portal.

4.2.2 Importação Simple XLS - Confirmação

É nesta fase que os produtos passarão das tabelas temporárias para as tabelas do portal.

Assim como todos os produtos já passaram pela verificação, não deve haver problemas na fase de transferência de dados de uma tabela para a outra. No entanto, se houver algum problema de rede ou do servidor em si, as tabelas temporárias e as tabelas do portal não ficam inconsistentes porque existe uma fase de transacionalidade [33] que começa no início da transferência e termina apenas com o *commit* dos dados depois de todas as tabelas do portal sejam preenchidas com sucesso.

Para o funcionamento correto, o portal utiliza sete tabelas diferentes nas quais guarda a informação dos produtos, ficheiros associados a eles e as suas traduções. Desta forma, é necessário transformar as 3 tabelas temporárias em 7 distintas. No entanto, na importação das

tabelas para o portal existem algumas verificações, como por exemplo a existência de um produto no portal (filtrando pela referência interna ou código de barras).

Para juntar a informação e guarda-la em tabelas do portal é utilizada a estrutura em dicionário do python que contém a informação necessária para cada tabela em questão. Esses dicionários vão ser processados pela biblioteca 4.2.7 Bulk. E assim acontece com todas as tabelas do portal.

A informação dos ficheiros também é guardada na base de dados e os ficheiros são codificados, junto com os thumbnails e transferidos para a aws Amazon [34] onde seguem a estrutura das pastas pelo ano, mês e dia. Os ficheiros são transferidos um a um garantindo dessa forma que não exista nenhum timeout do servidor e que o portal não fique subcarregado. Cada ficheiro ao ser carregado com sucesso é automaticamente removido da pasta temporária, pois já não vai ser utilizado mais no futuro.

No último passo é necessário atualizar as tabelas onde serão indexados todos detalhes do produto (nome, descrição, características e keyword) para que a pesquisa devolver resultados mais precisos. Nesta parte, durante o desenvolvimento, houve uma melhoria com qual conseguimos reduzir o tempo de carregamento de todos os produtos para metade. Como o carregamento inicial realizava 3 queries por cada produto foi decidido estudar melhor o código da biblioteca watson [35] de django e alterá-lo de forma a otimizar o seu resultado:

1. Em vez de guardar a informação do produto um a um decidiu-se fazer uma inserção em bulk de no máximo 100 produtos, reduzindo desta forma a execução de queries. Isto é, em 1000 produtos, em vez de executar 1000 queries passou a executar apenas 10.
2. Como no portal os produtos por norma são sempre inseridos e raramente modificados (modificação por norma é de poucos produtos e por isso é feita diretamente no portal) então é desnecessário seguir a política da biblioteca watson de estar a verificar se esse dado produto já existe na base de dados do portal. Isto porque se estamos a inserir produtos novos e já foi verificado que o produto não existe no portal, então tem de ser inserido um registo novo nas tabelas de indexação. Desta forma por inserção de um produto na base de dados poupamos uma query.

Na melhoria na biblioteca watson poupamos uma média de 2 queries por produto, o que reduz drasticamente o tempo de inserção dos produtos na base de dados.

Mais uma melhoria que foi efetuada durante o processo foi deixar de verificar o código de barras com os códigos já existentes no portal, caso este código seja um stamp gerado na fase de importação.

Depois de inserir todos os produtos no portal é verificado mais uma vez se todos os ficheiros ficaram removidos das pastas temporárias de ficheiros e é feita novamente a limpeza das tabelas temporárias do utilizador.

4.2.3 Importação Complex XLS - Carregamento temporário

A importação do Complex XLS começa por analisar o ficheiro xls e rejeitar a importação caso o ficheiro não siga as validações principais, tais como:

- Tamanho do ficheiro, maior que 0 Kb e menor que 30 Mb;
- Extensão do ficheiro tem de ser do tipo Excel: xls exlsx;
- 6 separadores no ficheiro e com os nomes corretos: ACTOR, CATALOG TREE, ITEM, FILE TO, SUPPLIER e PARAMETERS;
- Número de colunas de cada separador.

Depois de passar por essas validações serão avaliados todos os separadores um a um e registados os erros, caso estes existam.

Todos os registos importados com sucesso serão guardados em tabelas temporárias com uma estrutura simples, só uma tabela por separador, como se pode ver na Ilustração 3.

Os separadores são avaliados por lógica de construção duma empresa no portal. A primeira parte é inserção da empresa, contida no separador ACTOR.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	User Name	Cor banner	Nome legal	Short Name	País	NIF/VAT	Morada linha 1	Morada linha 2	Morada linha 3	Morada linha 4	Morada CEP	Localização GPS	Cod indicador telefonico	Telemóvel	Telefone	Fax	E-mail	URL
2	anton	#FFFFFF	Antonza Co	antonza	PT	247723436	Aveiro				3800 52.392609 1.270873		44	913129148	333577	562808	geral@ant.uk	ant.co.uk
3																		

Ilustração 20 - Ficheiro Complex XLS. Separador ACTOR - Preenchido

Este separador contém só uma linha de registo, como é exemplificado na Ilustração 20, e as outras linhas serão ignoradas. As validações têm de respeitar algumas regras para serem importadas para o portal. Caso estas regras não sejam respeitadas é registada a informação sobre o tipo de erro para ser apresentada ao utilizador (4.2.13 Tratamento de erros), as regras que estão indicadas nos requisitos do complex XLS.

Depois de verificar todos os campos do separador ACTOR, com ou sem sucesso, é transitado para o separador seguinte CATALOG TREE.

Este separador contém a informação do catálogo da empresa que está registada linha a linha e em que cada linha representa o nome do catálogo/sub-catálogo como mostra Ilustração 21.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6		Número do catálogo	Dependência do catálogo	Nome do catálogo	Título do catálogo	Informação do catálogo
2	Empresa							1	0	Empresa	Empresa	Empresa
3	About							2	0	About	About	About
4	Sobre							3	0	Sobre	Sobre	Sobre
5	Stratasys 3D Printer							4	0	Stratasys 3D Printer	Stratasys 3D Printer	Stratasys 3D Printer
6	About	About lv2_1						5	2	About lv2_1	About lv2_1	About lv2_1
7	About	About lv2_2						6	2	About lv2_2	About lv2_2	About lv2_2
8	About	About lv2_1	About lv3_1					7	5	About lv3_1	About lv3_1	About lv3_1
9	About	About lv2_1	About lv3_2					8	5	About lv3_2	About lv3_2	About lv3_2
10	About	About lv2_1	About lv3_3					9	5	About lv3_3	About lv3_3	About lv3_3
11	About	About lv2_1	About lv3_4					10	5	About lv3_4	About lv3_4	About lv3_4
12	Sobre	Sobre lv2_1						11	3	Sobre lv2_1	Sobre lv2_1	Sobre lv2_1
13	Sobre	Sobre lv2_2						12	3	Sobre lv2_2	Sobre lv2_2	Sobre lv2_2
14	Sobre	Sobre lv2_3						13	3	Sobre lv2_3	Sobre lv2_3	Sobre lv2_3
15	Sobre	Sobre lv2_4						14	3	Sobre lv2_4	Sobre lv2_4	Sobre lv2_4
16	Sobre	Sobre lv2_5						15	3	Sobre lv2_5	Sobre lv2_5	Sobre lv2_5
17	Sobre	Sobre lv2_6						16	3	Sobre lv2_6	Sobre lv2_6	Sobre lv2_6
18	Sobre	Sobre lv2_7						17	3	Sobre lv2_7	Sobre lv2_7	Sobre lv2_7
19												
20												

Ilustração 21 - Ficheiro Complex XLS. Separador CATALOG TREE - Preenchido

Durante a validação, todos os erros encontrados serão guardados para serem apresentados ao utilizador. Se existir pelo menos um erro neste separador então toda a estrutura desse catálogo não será transferida. Com ou sem erros o processo segue para o próximo separador ITEM.

Este separador contém a informação de produtos que vão ser inseridos no portal, cada linha do separador corresponde a um produto como mostra a Ilustração 22.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Ação	UPC do ITEM (se existir)	Referência interna actual ITEM	Nova referência	Nome ITEM	Características do ITEM	Especificações do ITEM	Papel da empresa nest ITEM (ACTOR ROLE)	Nº do CATALOG	Referência do fabricante para este ITEM	Nº/VAT fabricante deste ITEM	Nome do fabricante deste ITEM	URL do fabricante deste ITEM
2	ADD		ITEM1		ITEM1	Características 1	Especificações		2	11			
3	ADD		ITEM2		ITEM2	Características 2	Especificações		2	11			
4	ADD		ITEM3		ITEM3	Características 3	Especificações		2	11			
5	ADD		ITEM4		ITEM4	Características 4	Especificações		2	12	ITEM0102	Products	produto1.com
6	ADD		ITEM5		ITEM5	Características 5	Especificações		2	12			
7	ADD		ITEM6		ITEM6	Características 6	Especificações		2	12			
8	ADD		ITEM7		ITEM7	Características 7	Especificações		2	13			
9	ADD		ITEM8		ITEM8	Características 8	Especificações		2	13			
10	ADD		ITEM9		ITEM9	Características 9	Especificações		2	13	ITEM1423	Empresas	empresa.pt
11	ADD		ITEM10		ITEM10	Características 10	Especificações		2	4			
12	ADD		ITEM11		ITEM11	Características 11			2	4			
13	ADD		ITEM12		ITEM12	Características 12			2	4			
14	ADD		ITEM13		ITEM13	Características 13			2	4			
15	ADD		ITEM14		ITEM14	Características 14			2	23			
16	ADD		ITEM15		ITEM15	Características 15			2	23			
17	ADD		ITEM16		ITEM16	Características 16			2	23			
18	ADD		ITEM17		ITEM17	Características 17			2	23			
19	ADD		ITEM18		ITEM18	Características 18	Especificações		2	18			
20	ADD		ITEM19		ITEM19	Características 19	Especificações		2	18			
21	ADD		ITEM20		ITEM20	Características 20	Especificações		2	22			
22	ADD		ITEM21		ITEM21	Características 21	Especificações		2	22			
23	ADD		ITEM22		ITEM22	Características 22	Especificações		2	22			
24	ADD		ITEM23		ITEM23	Características 23			2	22			
25	ADD		ITEM24		ITEM24	Características 24			2	20			
26	ADD		ITEM25		ITEM25	Características 25			2	20			
27													
28													

Ilustração 22 - Ficheiro Complex XLS. Separador ITEM - Preenchido

Durante a validação, todos os erros encontrados, que estão indicadas nos requisitos do complex XLS, serão guardados para serem apresentados ao utilizador. Com ou sem erros o processo segue para o próximo separador FILE TO.

Este separador contém a informação de todos os ficheiros que serão associados a produtos, linhas de catálogo ou o logotipo de empresa. Cada linha deste separador corresponde a um ficheiro, como se mostra na Ilustração 23.

	A	B	C	D	E	F	G
	Para (ACTOR, CATALOG ou ITEM)	Tipo de ficheiro - Imagem ou PDF	Nome do Ficheiro	Destino (ref ITEM, nº CATALOG, shortname do ACTOR)	Nº de ordem dentro do mesmo destino	Alt / Título	Tooltip
1							
2	CATALOG	PDF	File01		14	1 ALT-Título 01	ALT-Título 01
3	CATALOG	PDF	File02		15	1 ALT-Título 02	ALT-Título 02
4	CATALOG	PDF	File03		15	2 ALT-Título 03	ALT-Título 03
5	CATALOG	IMG	File04		16	1 ALT-Título 04	ALT-Título 04
6	CATALOG	IMG	File05		16	2 ALT-Título 05	ALT-Título 05
7	CATALOG	PDF	File06		16	3 ALT-Título 06	ALT-Título 06
8	ACTOR	IMG	File07	ITEM1		1 ALT-Título 07	ALT-Título 07
9	ITEM	IMG	File08	ITEM2		1 ALT-Título 08	ALT-Título 08
10	ITEM	IMG	File09	ITEM3		1 ALT-Título 09	ALT-Título 09
11	ITEM	IMG	File10	ITEM4		1 ALT-Título 10	ALT-Título 10
12	ITEM	IMG	File11	ITEM5		1 ALT-Título 11	ALT-Título 11
13	ITEM	IMG	File12	ITEM6		1 ALT-Título 12	ALT-Título 12
14	ITEM	IMG	File22	ITEM15		2 ALT-Título 22	ALT-Título 22
15	ITEM	IMG	File23	ITEM17		1 ALT-Título 23	ALT-Título 23
16	ITEM	IMG	File27	ITEM20		2 ALT-Título 27	ALT-Título 27
17	ITEM	IMG	File28	ITEM22		1 ALT-Título 28	ALT-Título 28
18	ITEM	IMG	File29	ITEM23		1 ALT-Título 29	ALT-Título 29
19	ITEM	IMG	File30	ITEM23		2 ALT-Título 30	ALT-Título 30
20	CATALOG	PDF	File31		17	1 ALT-Título 31	ALT-Título 31
21	CATALOG	IMG	File32		17	1 ALT-Título 32	ALT-Título 32
22	CATALOG	IMG	File33		17	2 ALT-Título 33	ALT-Título 33

Ilustração 23 - Ficheiro Complex XLS. Separador FILE TO - Preenchido

Durante o processo de validação, todos os erros encontrados, que estão indicadas nos requisitos do complex XLS, serão guardados para serem apresentados ao utilizador. Com ou sem erros o processo segue para o próximo separador SUPPLIERS.

Este separador contém toda a informação sobre o fornecedor dos produtos. Cada linha do separador corresponde a um fornecedor de um produto, como mostra a Ilustração 24.

	A	B	C	D	E	F	
1	Referência interna ITEM	Referência do fornecedor	NIF/VAT do fornecedor	Nome do fornecedor			
2	ITEM3	ITEM3		Empresas			
3	ITEM4	ITEM4		Empresas			
4	ITEM5	ITEM5		Produtos			
5							

Ilustração 24 - Ficheiro Complex XLS. Separador SUPPLIERS - Preenchido

Depois de validar todos os registos de fornecedores, que estão indicadas nos requisitos do complex XLS, com ou sem erros na validação passamos para o último separador PARAMETERS.

Neste separador só existe uma linha e duas colunas, como mostra a Ilustração 25.

	A	B	C	D	E	F	G	
1	Idioma do Catálogo	ACTOR ID						
2	EN							

Atalhos: ACTOR CATALOG_TREE ITEM FILE TO SUPPLIERS PARAMETERS AJUDA

Ilustração 25 - Ficheiro Complex XLS. Separador PARAMETERS - Preenchido

Depois de validar todas as regras deste último separador, que estão indicadas nos requisitos do complex XLS, o programa irá reunir todos os tipos de erros encontrados no ficheiro xls e apresenta-los ao utilizador. Caso o ficheiro contenha mais de 11 erros, então não será possível confirmar e importar os produtos para o portal. Caso contrário o utilizador pode confirmar se tudo que foi inserido no ficheiro xls está correto.

Se existir ficheiros para associar, antes de confirmar a importação de produtos para o portal é necessário carregar todos os ficheiros que foram indicados no separador FILE TO. Este procedimento será descrito mais em baixo, no separador 4.2.8 Importação ficheiros.

Depois destes processos o utilizador pode confirmar a importação, de forma aos produtos que foram inseridos sejam transferidos de tabelas temporárias para o portal. Pode também abortar a importação. Ao abortar, todas as tabelas temporárias que foram preenchidos com os dados do ficheiro xls, serão limpas e todos os ficheiros para associar, caso já estejam carregados, serão removidos da pasta temporária.

4.2.4 Importação Complex XLS - Confirmação

Quando o utilizador confirmou todos os dados introduzidos e carregou no botão de confirmação será feito a transformação de registos das tabelas temporárias para uma estrutura de tabelas do portal. A estrutura do portal contem 16 tabelas para guardar a informação relativamente a empresa, catalogo, produtos, ficheiros, etc.

A transferência dos registos é feita registo a registo criando um dicionário com todos os registos a serem guardados e passado pelo 4.2.7 Bulk para a inserção rápida nas tabelas da base de dados do portal. Tal como no Simple XLS na inserção pelo Bulk criado de raiz, é necessária para garantir a transacionalidade [33] de informação.

Os ficheiros que estavam guardados na pasta temporária do utilizador serão enviados para a aws amazon e guardado o URL delas na base de dados do portal. Depois da importação de cada ficheiro este será removido da pasta temporária do utilizador.

Ao finalizar a transferência de registos para as tabelas do portal é necessário limpar todos os registos que existiam nas tabelas temporárias do utilizador que executava a importação.

4.2.5 Edição de produtos via XLS

A edição dos produtos é feita da forma semelhante a inserção. A diferença está na construção do ficheiro Excel.

Para o Complex XLS já tinha falado na coluna ACTION no separado ITEM onde é possível escolher a operação que tem de ser feita sobre o produto dessa linha.

4.2.6 Remoção de produtos via XLS

Para a remoção dos produtos do portal existe uma ferramenta no portal que permite uma seleção fácil dos produtos a serem removidos.

A remoção de produtos também é permitida pelo Complex XLS escolhendo a opção DELETE no separador ITEM na primeira coluna. Todos os itens selecionados irão ser removidos pela referência interna do produto. Caso os produtos identificados para remover não existirem é apresentado erro ao utilizador. No mesmo ficheiro Complex XLS é possível remover, adicionar ou modificar os produtos.

4.2.7 Bulk

A ferramenta bulk foi desenvolvida para importação de produtos do Excel para as tabelas temporárias e das tabelas temporárias para as tabelas do portal.

O Django já possuía uma biblioteca que inserisse os dados em bulk [36], mas durante o transporte dos produtos das tabelas temporárias para as tabelas do portal necessitamos de conhecer os ID's (primary key) de todos os produtos inseridos para usa-los como foreign key nas tabelas que necessitam da referencia do produtos. No entanto o bulk de Django não retornava os ID's dos produtos inseridos e como estamos a usar a transacionalidade então necessitamos das chaves dos registos já inseridos. Por isso foi desenvolvido a própria biblioteca bulk, que recebe como parâmetros o nome da tabela em qual vão ser inseridos os dados e um dicionário com informação a ser inserida, de seguinte forma {"nome": "nome do produto"}. A biblioteca ao receber os parâmetros de entrada cria uma query bulk, com o retorno dos id's inseridos, de postgresql e pela biblioteca do Django *connection cursor* [37] é enviada para a base de dados retornando os id's dos registos que foram inseridos. Esses id's vão ser necessários para poder utilizá-los ao inserir os outros registos em base de dados.

A velocidade de inserção pela biblioteca do Django e pela biblioteca criada de raiz é idêntica com a única melhoria que nos devolve a lista de chaves primárias dos registos inseridos.

4.2.8 Importação ficheiros

A importação dos ficheiros existe em duas partes distintas do portal, sendo uma delas na importação do ficheiro XLS e outra na importação exclusiva dos ficheiros para o repositório do utilizador. Importar os ficheiros para o repositório do utilizador faz com que ele possa associar os ficheiros aos seus produtos utilizando ferramentas próprias do portal.

A lógica das importações é semelhante, a única diferença é que num dos casos os ficheiros ficam logo associados aos produtos e só serão importados aqueles ficheiros que estão ligados aos produtos no ficheiro XLS. Ao importar os ficheiros para o repositório não existe qualquer restrição em relação a quantidade de ficheiros a serem importados.

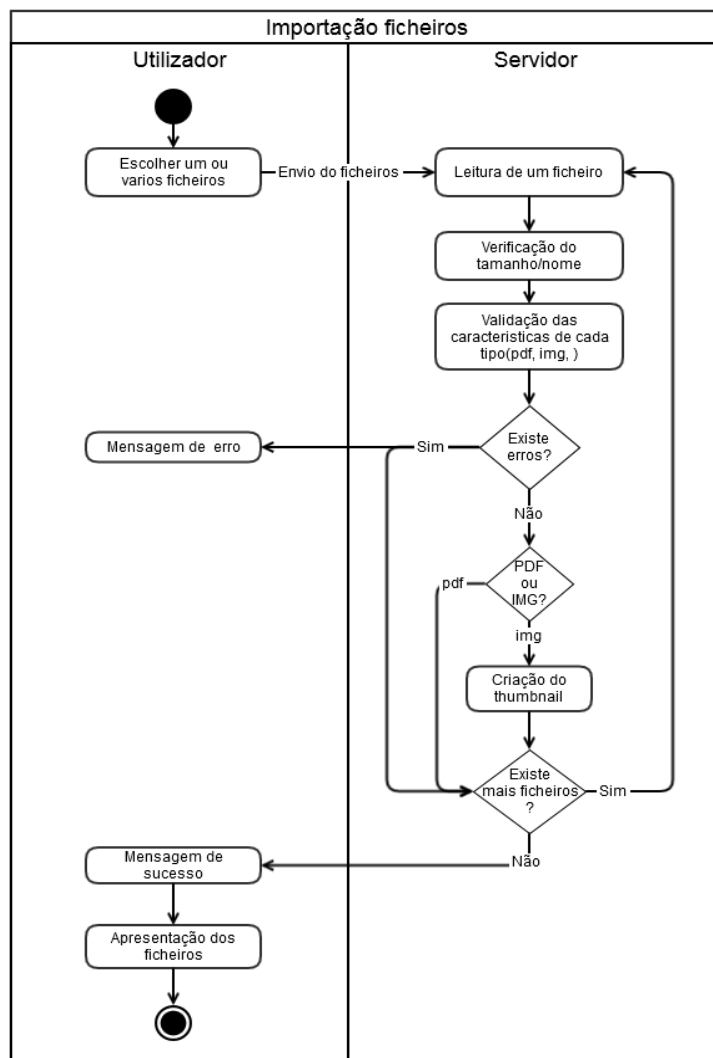


Ilustração 26 - Diagrama de atividade - Importação de ficheiros

A importação dos ficheiros utiliza a form de Django que importa o ficheiro e guarda-o na memória para ser tratado. Os ficheiros podem ser selecionados em massa, mas a verificação será feita um a um para descartar ficheiros com problemas e assim poupar no processamento. A apresentação breve está na Ilustração 26. As regras de importação encontram-se nas Tabela 11, Tabela 12 e Tabela 13. Cada ficheiro passa por um conjunto de operações de validação, que são:

- Extensão (.pdf, .jpg, e .jpeg) e tipo do ficheiro. Por vezes só a extensão não chega para garantir que o ficheiro seja dum determinado tipo. Para isso ocorreremos a algumas ferramentas para descobrir isso:

- Para PDF's utilizamos uma biblioteca python pdfminer/PDFParser[38] que simula a abertura do ficheiro PDF e no caso do sucesso garantimos que o ficheiro que está a ser importado é mesmo do tipo pdf;
- Para imagens também é usada uma ferramenta, a biblioteca PIL [39] que nos indica se o ficheiro que está a ser importado é do tipo imagem;
- O tamanho do ficheiro não pode ultrapassar os 3 Mb, isto foi definido para não subcarregar o portal, visto que cada produto pode ter múltiplas imagens e pdf's;
- O tamanho do nome do ficheiro não pode ultrapassar os 40 caracteres, isso serve para garantir a coerência dos nomes no servidor onde serão guardadas as imagens.
- Validar se o ficheiro a ser importado ainda não existe na base de dados do próprio utilizador, caso exista a imagem será descartada e a imagem já existente é reutilizada para o produto.
- Validar se o ficheiro a ser importado foi descrito no ficheiro xls, preenchido pelo utilizador. Caso o ficheiro que está a ser transferido não existira no ficheiro xls então é descartado porque não há onde é possível associa-lo.
- Validação dos tamanhos e proporção das imagens.

Cada ficheiro do tipo imagem ainda é tratado para diminuir o tamanho do ficheiro original e é criado um thumbnail para ser usado no portal.

Todos os ficheiros importados serão guardados numa pasta Media com os nomes cifrados [40] para que na base dados do portal não serem reconhecidos pelas entidades externas.

No caso do erro em pelo menos um dos pontos, o ficheiro não é aceite e é apresentada uma mensagem de erro ao utilizador. Caso existam ficheiros com o mesmo nome um deles será descartado e não apresentado ao utilizador.

Os nomes originais do ficheiro e os nomes codificados serão guardados na sessão do utilizador para serem confrontados com os nomes dos ficheiros inseridos pelo utilizador através do xls.

Durante a inserção de um conjunto de ficheiros é apresentada uma barra de conclusão ao utilizador que dá uma imagem do estado em que está uma determinada importação de ficheiros. Quando a barra verde chegar ao fim, isso quer dizer que todos os ficheiros foram transferidos para o portal, sendo eles com (Ilustração 28) ou sem erros (Ilustração 27).

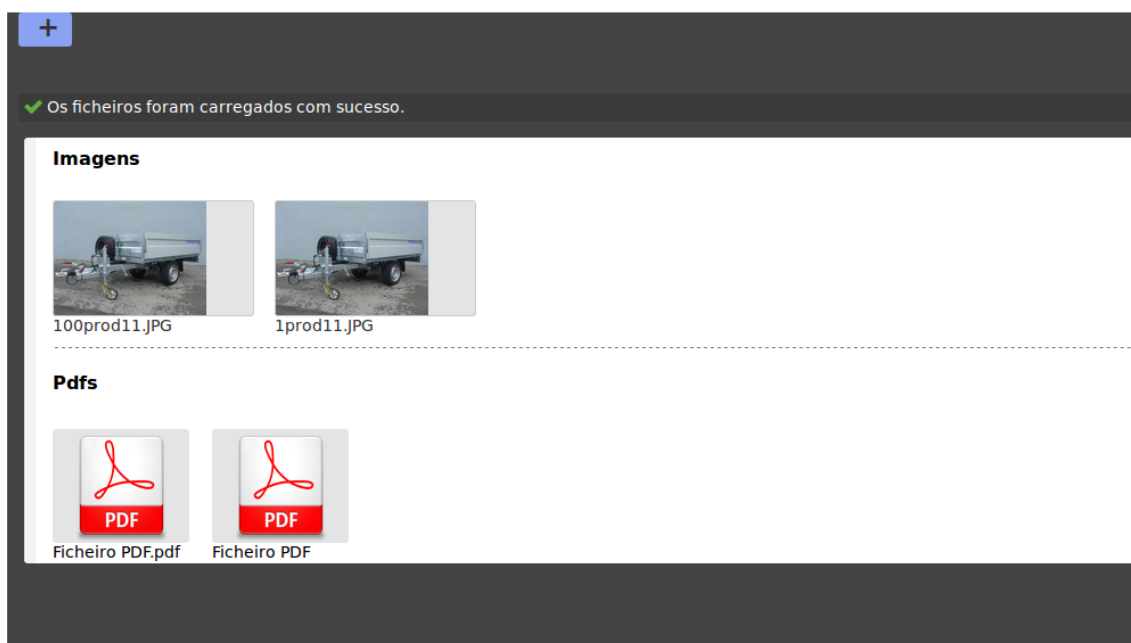


Ilustração 27 - Importação de ficheiros - Sem erros



Ilustração 28 - Importação de ficheiros - Com erros

Condições para os ficheiros de associação:

Nome	Condições	
Tipo	.jpg .jpeg	
	Mínimo	Máximo
Tamanho do ficheiro	> 0Mb	3Mb
Tamanho do nome	1 caracter	40 caracteres
Largura da imagem	200 pixeis	Sem limite
Altura da imagem	200 pixeis	Sem limite
Requisitos	Imagem válida	
Proporção	1:5	

Caracteres inválidos	/\ ; "
-----------------------------	--------

Tabela 11 - Verificações dos ficheiros do tipo imagem

Nome	Condições	
Tipo	.png	
	Mínimo	Máximo
Tamanho do ficheiro	> 0Mb	3Mb
Tamanho do nome	1 caracter	40 caracteres
Largura da imagem	156 pixeis	Sem limite
Altura da imagem	64 pixeis	Sem limite
Requisitos	Imagem válida, Largura > Altura	
Proporção	1:2,4	
Caracteres inválidos	/\ ; "	

Tabela 12 - Verificação dos ficheiros do tipo imagem

Nome	Condições	
Tipo	.pdf	
	Mínimo	Máximo
Tamanho do ficheiro	> 0Mb	3Mb
Tamanho do nome	1 caracter	40 caracteres
Requisitos	PDF válido	
Caracteres inválidos	/\ ; "	

Tabela 13 - Verificações dos ficheiros do tipo pdf

4.2.9 Importação ficheiro ZIP

Ao longo das importações de produtos das diversas empresas foi decidido criar algo que facilite o trabalho da importação e melhore a performance. A ferramenta ZIP [41] ajudou bastante nesta situação. A importação através do ficheiro ZIP é possível com os ficheiros Simple XLS e Complex XLS.

Para criar um ficheiro ZIP é necessário criar os ficheiros XLS (4.1.1 Simple XLS (simples) e 4.1.2 Complex XLS(simples)), escolher todos os ficheiros que foram indicados no ficheiro XLS e incluir isso tudo no ficheiro ZIP. As imagens e os ficheiros PDF podem ser separados por pastas diferentes para facilitar a organização dos ficheiros.

Desta forma o ficheiro XLS e o conjunto dos ficheiros associados aos produtos ficam mais comprimidos e são transferidos mais rapidamente para o servidor. Desde que não sejam demasiados ficheiros a associar, quanto maior for o ficheiro ZIP mais complexo é para processa-lo no servidor.

Para ler o ficheiro ZIP utilizei a biblioteca python zipfile [42]. Esta biblioteca, extrai todos os ficheiros para uma pasta temporária, e começa o processamento de todos os produtos existentes

no ficheiro XLS e verificação de todas as imagens e pdf's. Caso esteja tudo correto é apresentado o resultado com os produtos do ficheiro XLS e os ficheiros associados a cada produto.

A parte de confirmação segue exatamente a mesma estrutura do ponto 4.2.2 Importação Simple XLS - Confirmação ou 4.2.4 Importação Complex XLS - Confirmação.

Desta forma a importação segue as mesmas regras e regista de forma igual, como se estivéssemos a importar ficheiros xls e ficheiros para associar separadamente.

4.2.10 Multilingue

A multilingue está presente em todo o portal, desde dos botões até a descrição de erros na importação dos ficheiros XLS. Desta forma ao importar os produtos para o portal, por exemplo, e ocorram erros durante a importação, o utilizador facilmente consegue perceber o que está errado, porque o tipo de erro é lhe apresentado numa língua que ele consiga entender.

A tradução do portal está feita em 7 línguas diferentes, no entanto os produtos podem ser importados e guardados no portal em qualquer língua, desde que o utilizador saiba preencher corretamente a informação sobre o produto.

Caso o utilizador só conheça uma das línguas, existe também a possibilidade de usar tradutor (2.6.3 Google Translate API) automático integrado na importação dos produtos. Neste caso o utilizador antes de importar o ficheiro terá que escolher para que línguas pretende traduzir os produtos que estão no ficheiro XLS. Desta forma o utilizador ao fazer importação dos seus produtos numa língua, automaticamente irá ter os produtos disponíveis nas línguas por ele escolhidas, isto serve para ter o maior público-alvo e facilitará as vendas dos produtos.

A performance neste caso pode ser mais baixa, dependendo para quantas línguas é feita a tradução e dependendo da velocidade da internet e da API translate, uma vez que por cada célula de informação do produto que está no ficheiro XLS é necessário fazer um pedido de tradução para uma ou mais línguas.

Todo este processo foi desenvolvido para garantir que todos os utilizadores possam transferir produtos em todas as línguas que eles estiverem interessados para poder obter um número maior de clientes.

4.2.11 GPC categorias

As categorias [29] são utilizadas para guardar os produtos da forma mais adequada no portal. Desta forma é mais fácil de encontrar uma determinada categoria com os produtos existentes nela.

A GPC é um standard das categorias de produtos organizado numa árvore para as separar de uma forma mais precisa. As categorias estão disponibilizadas em formato XLS ou XML e seguem a seguinte estrutura: Segmento, Family, Class, Brick, Attribute e AttributeValue. Um exemplo de

uma determinada categoria e a especificação do produto é: Alimentos / Bebidas / Tabaco > Pão / Panificação > Pão > Pães (Congelados). Isto é uma sequência de categorias para chegar aos atributos específicos, para o nosso caso temos atributos e os valores deles: Tipo de Pão - BAGEL|BAGUETTE|BATARD|etc., Tipo de cereal/grão - BRAN|CORN|MAIZE|etc.

Todas as categorias em várias línguas são importadas (4.3.2 Importação de categorias GPC) para o portal para os utilizadores as poderem utilizar e desta forma permitir organizar os seus produtos em categorias corretas. As categorias podem ser escolhidas pelo utilizador durante o download do ficheiro XLS ou utiliza-las quando os produtos já estejam inseridos no portal.

4.2.12 Integração com Magento

Para poder importar os produtos para o Magento é necessário criar uma conta e dar permissões de administrador. Depois é necessário gerar uma key de API para poder fazer operações sobre os produtos que estão já na plataforma Magento e aqueles que vão ser transferidos.

Para importar os produtos para o Magento o utilizador tem de escolher todos os produtos que pretende importar e clicar no botão de envio.

Para a importação utilizamos o XML-RPC cliente [43] que tem os métodos que nos interessa. No nosso caso precisamos de importar os produtos e imagens do nosso portal para a plataforma Magento. Para o estabelecimento de comunicação executamos os passos do Código 6.

```
def products(request):
    username = "antonza"
    password = "0ced2e13bcdd751b3d6ee30f0b187443"
    server =
    apiserwer('http://localhost/magento/index.php/api/xmlrpc/')
    session = apisession(server, username, password)

def apiserwer(address):
    """connect to server return server object"""
    return xmlrpclib.Server(address)

def apisession(server, username, password):
    """open an api session and return the string"""
    return server.login(username, password)
```

Código 6 - Ligação com API XML-RPC

A primeira operação a fazer é criação de produto na plataforma Magento. Para isso usamos o método `catalog_product.create` [44] que recebe a informação sobre o produto no formato JSON. O ficheiro JSON é preenchido com a informação de cada produto, mas existe campos obrigatórios sem quais a importação fica sem efeito, entre os quais, o nome do produto, a referência, estado visível, etc. É possível também indicar alguns campos adicionais como a existência de produtos em stock, o preço deles, etc., como mostra o Código 7.

```

def importProduct(request, server, session, itens):
    """create a product using the api"""
    productSel = request.POST.getlist('product_check_list')
    itens = itens.filter(id__in = productSel)

    # retrieve current set information
    currset = server.call(session, 'product_attribute_set.list')
    currsetid = currset[0].get('set_id')
    for item in itens:
        stockData = {
            'use_config_manage_stock':1,
            'manage_stock':1,
            'is_in_stock':1,
            'qty':"20",
        }

        proddetail = dict({
            'name':item.item_name,
            'websites':[1],
            'short_description':item.item_features,
            'description':item.item_tech_spec,
            'price':"0",
            'weight':"0",
            'categories':['3', '6', '17'],
            'status':1,
            'tax_class_id':"2",
        })
        newprod = ["simple", currsetid, prodsku, proddetail]

    try:
        result = server.call(session, 'product.create', newprod)
    except:
        print "Falha na comunicação. Tente outra vez"

```

Código 7 - Criação de produto em Magento

A segunda operação a fazer será a importação de imagens que deverão estar associadas aos produtos. Isto permite fazer com que toda a informação de produtos que existe no portal seja igual no Magento. Para isso utilizamos o método `product_media.create` [45] que serve para a criação de novas imagens e associá-las a uma determinada referência. Da mesma forma, indicamos que as imagens importadas têm de servir como imagem principal do produto e como thumbnail da imagem. Para execução da operação utilizamos o Código 8.

```

def createImage(self, sku, itemActorId, srv, session):
    files = File.objects.select_related().filter(id = Id)
    for reg in files:
        if(reg.file.mime_type == "image/jpeg"):
            link = 'https://portal.s3.aws.com/' + reg.file.loc
            image = urllib.urlopen(link)
            try:
                encoded_string = base64.encodestring(image.read())
                img = {
                    'file':{
                        'content':encoded_string,
                        'mime':'image/jpeg',
                    },
                    'label':reg.file.file_name,
                    'position':i,
                    'types':["small_image", "thumbnail", "image"],
                    'exclude':0,
                }
                srv.call(session,'product_media.create',[sku, img])
            except:
                print "erro ao inserir a imagem: " + link

```

Código 8 - Criação de uma imagem em Magento

A informação dos produtos importados será guardada na base de dados do Magento. Por defeito a base de dados é MySQL [46], com estrutura já construída pela plataforma Magento. Depois da importação dos produtos, os mesmos ficam disponibilizados na plataforma Magento. As imagens também são guardadas em MySQL previamente codificadas em base64 através de biblioteca python base64 [47].

A importação dos produtos para a plataforma demora algum tempo por causa da estrutura complexa do Magento e pela codificação das imagens. Para evitar que a plataforma fique presa ao fazer a importação, este processo é posto em background, Código 9.

```

import threading
#call function in background
bground = threading.Thread(target=magento.importProduct, args = ())
#define function and run in background
bground.daemon = True
#start running
bground.start()

```

Código 9 - Execução processos em background

4.2.13 Tratamento de erros

O tratamento de erros é uma parte muito essencial na importação dos produtos, porque é com os erros que o utilizador consegue perceber o que está mal no ficheiro XLS para poder corrigir e importar novamente com sucesso.

Os erros são disponíveis em todas as línguas do site (4.2.10 Multilingue) para satisfazer as necessidades de qualquer utilizador que está a utilizar o portal. A apresentação do erro vem na língua em que o utilizador está a ver o portal, mas também pode haver casos em que o utilizador

tem o site numa língua em que não exista a tradução do erro, neste caso é apresentada a língua que vem por defeito no portal, inglês.

A lista de erros está guardada num dicionário interno no servidor para que o acesso seja mais rápido e não existir acessos desnecessários a base de dados. Os dicionários contêm a seguinte estrutura: {Lingua:{variavelErro:descriçãoErro}} -> {'PT': {'outSize2': 'Tamanho incorreto. Tem de ter 2 caracteres.'}}. Com essa estrutura conseguimos guardar todos os erros em todas as línguas e utiliza-los sempre que haja um determinado erro ao importar o ficheiro XLS ou outros ficheiros do tipo imagem e pdf.

No total existe descrição de 133 erros diferentes que podem aparecer ao importar os ficheiros XLS ou os ficheiros que vão ser associados aos produtos.

Durante a importação do ficheiro XLS na fase de pré-carregamento os erros são guardados numa tabela de base de dados com os seguintes parâmetros: Nome do utilizador, nome do ficheiro, tamanho do ficheiro, separador onde aconteceu o erro, a célula exata do erro, variável do erro, data e hora do erro e o identificador da importação. Estes dados todos são necessários para, na lista de erros, escolher a descrição correta do erro, a posição exata e apresentar ao utilizador que esta a fazer a importação (ex.: Linha 3 Coluna B. Separador FILE). O resto dos dados que estão na tabela de erros são necessários para, em trabalho futuro, possamos descobrir e corrigir os erros que os utilizadores mais cometem.

O utilizador não consegue passar para a fase seguinte, fase de confirmação se existir mais de 11 erros no ficheiro xls. Se o ficheiro apresentar menos de 11 erros é possível confirmar a importação caso o utilizador pretenda. Um exemplo da importação com alguns erros é apresentado na Ilustração 29

As cores apresentadas são dependentes da quantidade de erros:

- 0 erros, a cor apresentada é verde e tem uma mensagem de sucesso;
- Menos de 11 erros, a cor apresentada é laranja e com uma mensagem a avisar o utilizador e que ao confirmar nem todos os produtos serão importados para o portal;
- Mais de 11 erros, a cor apresentada é vermelha e é apresentada uma mensagem ao utilizador que não é possível importar os produtos e que é necessário corrigir os erros e tentar novamente.

CPImportXLS_10001426.xlsx

File size:

7859 kb

File language:

pt

File status:

This file has more than 11 errors. Please correct the file and try again.

Verify the XLS file and try again.

Error no.:

23

View details report

Details upload report

N	Sheet	Error position	Error description
1	ITEM	Line 5 Column A	This cell must have 2-50 characters.
2	ITEM	Line 5 Column E	This is a required cell.
3	ITEM	Line 6 Column A	This cell must have 2-50 characters.
4	ITEM	Line 6 Column A, B, C	This is a required cell.
5	ITEM	Line 6 Column F	The file name can only have 100 characters maximum.
6	ITEM	Line 6 Column D	Verify the barcode: 3232
7	ITEM	Line 7 Column A	This cell must have 2-50 characters.
8	ITEM	Line 7 Column A, B, C, E	This is a required cell.
9	ITEM	Line 7 Column D	Verify the barcode: casc
10	ITEM	Line 8 Column A	This cell must have 2-50 characters.
11	ITEM	Line 8 Column A, B, C, E	This is a required cell.
12	FILE TO	Line 2 Column C	This product does not exist in this XLS file.
13	FILE TO	Line 2 Column E	This cell only allows 30 characters maximum.
14	FILE TO	Line 2 Column A	This is a required cell.
15	FILE TO	Line 2 Column A	Choose one of the following types: PDF or IMG.

Ilustração 29 - Importação XLS - Apresentação de erros

Na importação de ficheiros para o repositório já não existe o mecanismo de guardar os erros numa base de dados, visto que estes não são necessários para uma utilização futura. Ao importar um ficheiro do tipo imagem ou pdf e caso o ficheiro não conseguir passar por todas as verificações o erro é dado em tempo real, ao detetar em que fase o erro foi descoberto. Assim, no momento do erro é feito um acesso ao dicionário dos erros e recolhida a descrição do erro na língua do site do utilizador. O erro é apresentado em vermelho, como apresenta a Ilustração 30.

Select and upload files

Browse....

No files selected.

Accepted types: PDF, JPG and JPEG

Error inserting file.
Only PDF, JPG, JPEG files are allowed. Please verify the following files: xgoogle.zip

Error inserting file.
The file name must have 40 characters maximum. Please verify the following files: NpmeDo PDFNpmeDo PDFNpmeDo PDFNpmeDo
PDFNpmeDo PDFNpmeDo PDFNpmeDo PDFNpmeDo PDFNpmeDo PDFNpmeDo PDF.pdf

Error inserting file.
Only PDF, JPG, JPEG files are allowed. Please verify the following files: h9sn02ww.exe

Error inserting file.
Only PDF, JPG, JPEG files are allowed. Please verify the following files: CPIImportXLS_10001426.xlsx

Error inserting file.
Invalid image. The image must be at least 200px x 200px and its proportion must be smaller than 1:5.CPIImportXLS_10000785.jpg

File inserted successfully.

Ilustração 30 - Importação de ficheiro - Apresentação de erros

4.3 Outras importações/exportações

Existiu também a necessidade em criar diferentes tipos de importação de ficheiros xls, para facilitar a forma de trabalhar para o resto da equipa do portal.

Uma vez que a ferramenta Excel é muito fácil de ser manipulada e rápida em organização decidimos manter a importação de ficheiros com o formato xls.

Foram também desenvolvidas algumas ferramentas para exportação de dados, é uma parte fundamental para a notificação de utilizadores sobre os novos produtos e novas atualizações do portal.

Como por exemplo, com o MailChimp conseguimos exportar os dados de utilizadores para a plataforma MailChimp.

4.3.1 Importação de variáveis dinâmicas

A importação do ficheiro com as variáveis tem o principal objetivo de facilitar a organização e a sua tradução. As variáveis no portal são utilizadas em vários pontos, como por exemplo, botões, nomes das páginas, títulos, avisos durante a navegação, etc.

O ficheiro Excel a ser importado contém 12 separadores, em que cada um corresponde a diferentes temas do site. Cada separador tem várias colunas:

- A primeira é a coluna com o nome da variável que irá ser usada no código para apresentar um determinado campo (botão, aviso, etc.);
- A segunda coluna é uma breve descrição da variável;
- A terceira coluna contém o destino da variável, isto é, a variável pode ser atribuída à parte Desktop, Mobile ou ambos os destinos;
- As próximas 7 colunas são as colunas de diferentes línguas (PT, EN, RU, etc). Essas línguas seriam usadas para apresentar ao utilizador o portal em 7 línguas diferentes.

A importação do ficheiro xls para o portal efetuava a transformação de todos os campos para um ficheiro JSON. Posteriormente criava o cabeçalho necessário para ser utilizador no portal da forma correta e era atribuído um nome com a extensão .json para ser compatível para a leitura da ferramenta de front-end. Através do ficheiro do formato json as variáveis são lidas e apresentadas ao utilizador final de uma forma instantânea.

4.3.2 Importação de categorias GPC

A necessidade de importação de ficheiros GPC veio logo que este standard de categorias foi encontrado pelo gestor de marketing. O GPC disponibiliza todas as categorias de produtos [29] para usar livremente, essas categorias existem em dois formatos:

- XLS - o ficheiro é construído em inglês, o resto das línguas existem no outro formato, com 12 colunas sendo cada par de colunas indica o código e o nome da categoria do Segment, Family, Class, Brick, Atributo e valor desse atributo.
- XML - o ficheiro contém só uma linguagem e esta na estrutura apresentada no Código 10.

```
<?xml version="1.0" encoding="UTF-8"?><schema>
  <segment code="6800" text="Segment" definition="">
    <family code="6801" text="Family" definition="">
      <class code="6802" text="Class" definition="">
        <brick code="1000" text="Brick" definition="Des">
          <attType code="2000" text="Atributo" definition=" Desc">
            <attValue code="3000" text="Valor" definition=""/>
          </attType>
        </brick>
      </class>
    </family>
  </segment>
</schema>
```

Código 10 - GPC Categorias - Estrutura XML

Para importar todas as categorias em XLS utilizei a biblioteca xlrd [32] guardando todos os valores lidos num dicionário. De seguida, através do método 4.2.7 Bulk, são inseridas numa estrutura própria na base de dados do portal, onde é armazenado o código da categoria, o valor da categoria e a língua da categoria.

Para importar o ficheiro XML foi utilizada biblioteca xml.dom [48], a qual apresenta fácil utilização. As categorias lidas serão guardadas num dicionário e, através do método Bulk, guardadas numa estrutura própria na base de dados do portal. São armazenados todos os valores e também a descrição caso exista. No Código 11 está um exemplo de execução das operações realizadas com a biblioteca xml.dom.

```
from xml.dom import minidom
xmldoc = minidom.parse("/GS1.xml")
segmentList = xmldoc.getElementsByTagName('segment')
for reg in segmentList:
    print reg.attributes['code'].value + "|" +
          reg.attributes['text'].value + "|" +
          reg.attributes['definition'].value
```

Código 11 - GPC Categorias - Leitura XML

Os standards das categorias são por vezes atualizados pelo equipa GS1. Quando isso acontece é necessário atualizar os valores que já existem na base de dados e acrescentar novas categorias, caso estas existam. Para isso utilizamos o mesmo código que temos para a leitura mas desta vez com duas partes distintas:

- Leitura e comparação de categorias novas com as categorias e valores já existentes no portal. Extrair a diferença de valores, novos valores e valores que já não existem, mostrando isso ao utilizador para estes serem verificados;
- Caso o utilizador concordar com as alterações que vão ser feitas e que categorias novas vão ser importadas para o portal então pode confirmar a importação.

Depois de confirmar, novas categorias serão adicionadas ao portal e as existentes que sofreram alterações vão ser alteradas.

4.3.3 MailChimp

MailChimp é uma plataforma responsável pela distribuição de Email com diferentes destinos e com textos dinâmicos, 2.6.2 MailChimp.

A necessidade de utilizar esta plataforma veio logo que o número de clientes começou a crescer. Para mandar um Email de cada vez para diferentes utilizadores requeria muito trabalho e tempo, por isso o MailChimp oferece-nos todas as ferramentas necessárias para otimizar este processo. Para que as ferramentas funcionem é necessário exportar a lista de utilizadores e templates de email que pretendemos utilizar.

O primeiro passo a fazer é a criação da conta no MailChimp [18] e gerar uma API Key que faz a ligação entre o nosso portal e a plataforma MailChimp. Depois de obter a API Key é necessário fazer a ligação do nosso lado do portal.

A ligação com o portal pode ser feita por duas aplicações diferentes, cada uma tem as suas facilidades sobre a plataforma MailChimp. A primeira é a mailchimp [49], que é fornecida diretamente da plataforma MailChimp, com esta API é mais fácil a leitura de registos que temos atualmente na plataforma, envio de campanhas de email e obtenção de reports sobre os email enviados. A segunda API utilizada é o chimpy [50], uma API que adaptou e facilitou algumas operações efetuadas sobre a plataforma referida. Esta API facilita nos registar novos utilizadores na plataforma e ajuda a obter dados mais detalhados sobre o report. O exemplo de uma conexão a cada API está no Código 12.

```
from chimpy import chimpy
import mailchimp

def connectMC(key):
    mc = mailchimp.Mailchimp(key)
    return mc

def connectChimp(key):
    chimp = chimpy.Connection(key)
    return chimp
```

Código 12 - Conectar-se a API MailChimp

Depois da conexão ser estabelecida podemos efetuar diversas operações sobre a plataforma MailChimp, algumas das quais eu passo a explicar.

Utilizadores

Os utilizadores na plataforma MailChimp são essenciais para poder enviar um ou vários Emails. Para isso é necessário exportar os utilizadores existentes no nosso portal. Criamos uma página de administrador que retorna todos os utilizadores existentes no sistema com alguma informação sobre eles, como por exemplo primeiro nome, ultimo nome, nome das empresas, Email e URL das empresas. Estes dados serão então exportados para o MailChimp. O exemplo da exportação está presente no Código 13. Antes de efetuar as subscrições é obtida a lista de todos os utilizadores

existentes no MailChimp, para que não sejam exportados utilizadores já existentes na plataforma de envio de Emails.

```
# existed members
members = []
for i in chimp.list_members(mcList):
    members.append(i['email'])

for reg in Utilizadores:
    if(reg['EMAIL'] not in members):
        aux = {"FN":"Anton", "LN":"Zverev", "URL":"anton.pt"}
        # Subscribe to list
        try:
            chimp.list_subscribe(mcList, reg.mail, aux)
        except:
            print "não foi possível adicionar"
```

Código 13 - MailChimp - Exportação de utilizadores - Subscrições

Email

Depois de existir a lista de utilizadores no MailChimp é necessário que os utilizadores recebam os emails sobre a informação de novos produtos, promoções, etc. Para realizar isso é necessário criar um template com a estrutura do Email e indicar os campos dinâmicos que serão diferentes para os utilizadores. Para exportar o template do Email para o MailChimp foi criada uma pagina onde é possível inserir código HTML e estruturar o email da melhor forma. Os campos dinâmicos no código HTML são identificados de seguinte forma: `*|FNAME|*`, `*|LNAME|*`, que correspondem ao primeiro e ultimo nome, estes campos dinâmicos são geridos diretamente na plataforma MailChimp. Para o envio do template utilizamos o exemplo presente no Código 14.

```
# send html code to mailchimp
mc.templates.add("Nome Template", htmlCode)
```

Código 14 - MailChimp - Envio da template

Com uma lista de emails completa e com o template criado é necessário enviar a mensagem desejada para a lista de utilizadores. O envio pode ser efetuado no momento ou numa determinada data a definir, ambos os exemplos estão no Código 15.

```

campCreate = mc.campaigns.create("regular", {"list_id": listID,
"subject": subject, "from_email": fromMail, "from_name": sendName,
"template_id": tempID}, "")
campID = campCreate['id']
# for immediately send
if(calendar == None):
    # send created campaign
    send = mc.campaigns.send(campID)
# send with schudule
else:
    # Create correct datetime format
    datetimeX = str(calendar.date()) + ' 0:00:00'
    # schedule campaign in mailchimp
    send = mc.campaigns.schedule(campID, datetimeX)

```

Código 15 - MailChimp - Envio Emails

Relatório de envio

Depois do envio das mensagens para os utilizadores é possível visualizar o estado da entrega dos emails, aqueles que foram recebidos corretamente, os que não conseguem ser entregues e aquelas em que o servidor não esta online. Para importar os resultados para o nosso portal executamos operações sobre ambas as API's, sendo que uma complementa a outra. No Código 16 está o exemplo para obter os resultados de envio.

```

# get all report list from one campaign sending
sentTo = mc.reports.sent_to(idEmailCampaign) ['data']
softBounces = chimp.campaign_soft_bounces(idEmailCampaign)
hardBounces = chimp.campaign_hard_bounces(idEmailCampaign)

```

Código 16 - MailChimp - Importação de resultados

5. Resultados

Ao longo do projeto foram realizadas várias análises relativamente aos resultados de certas operações, nomeadamente no que diz respeito aos tempos de execução das operações import/export. Estes resultados potenciaram alterações em certas partes do código bem como a utilização de novas bibliotecas python que melhoraram de forma significativa o desempenho das operações de import/export. A escolha das bibliotecas utilizadas teve como fator decisivo a adaptabilidade das mesmas ao portal desenvolvido. A importação de produtos com os ficheiros associados para o portal é executada em duas fases: temporária (valores apresentados na Tabela 14) e de confirmação (valores apresentados na Tabela 15). Para este caso foram realizados um conjunto de teste para medir em detalhe a performance do Simple XLS.

Os testes foram executados para 1, 10, 100, 1000 e 10000 produtos e com os seguintes anexos:

1. 0 ficheiros associados ao produto;
2. Uma imagem associada a cada produto;
3. Duas imagens associadas a cada produto;
4. Três imagens associadas a cada produto;
5. Três imagens e um pdf associados a cada produto;
6. Três imagens e dois pdf's associados a cada produto.

Imagem de teste com +- 80Kb e pdf de teste com +- 2Mb.

Produtos\Ficheiros	0	1	2	3	4	5
1	0.1s	0.1s	0.2s	0.2s	0.4s	0.7s
10	0.4s	0.7s	1.1s	1.4s	3.8s	6.2s
100	0.4s	4.0s	7.6s	11.2s	34.9s	58.6s
1000	0.6s	36.6s	72.6s	108.6s	345.6s	582.6s
10000	16.6s	376.6s	736.6s	1096.6s	3466.6s	5836.6s
Tempo médio	0.00s	0.04s	0.07s	0.11s	0.35s	0.58s

Tabela 14 - Importação de produtos para o portal na fase de pré-carregamento

Produtos\Ficheiros	0	1	2	3	4	5
1	0.8s	3.5s	6.2s	8.9s	20.6s	32.3s
10	1.2s	28.1s	55.1s	82.0s	199.0s	316.0s
100	1.2s	270.4s	539.6s	808.8s	1978.8s	3148.8s
1000	2.1s	2694.1s	5386.1s	8078.1s	19778.1s	31478.1s
10000	24.3s	26944.3s	53864.3s	80784.3s	197784.3s	314784.3s
Tempo médio	0.00s	2.69s	5.39s	8.08s	19.78s	31.48s

Tabela 15 - Importação de produtos para o portal na fase de confirmação

Nota: Os resultados na fase de importação são demasiado elevados devido a velocidade da internet por parte do emissor e por parte do remetente.

Resultados:

1. A importação de produtos é muito rápida, sendo de um ou de 10000 produtos;
2. A performance começa a descer com acréscimo de produtos por importação, devido as verificações de cada produto importado.
3. A performance também desce com a quantidade de ficheiros associados a cada produto, devido a velocidade de internet de upload e capacidade de servidor receber os ficheiros. Este ponto pode ser melhorado caso o servidor do portal for mais rápido e a internet do utilizador for superior.

Durante muitas importações de produtos para o portal reparou-se que o utilizador é obrigado de passar por varias etapas(importar xls, imagens, pdf's) para importar os seus produtos com ficheiros associados para o portal, então decidiu-se em criar um diferente tipo de importação. Este tipo de importação é através de ficheiro ZIP.

Para que seja possível medir o desempenho de importação de produtos com e sem o ficheiro ZIP, os testes foram executados com o mesmo numero de produtos e ficheiros. Os resultados da Tabela 16 não apresentam importação de 10000 produtos com ficheiros associados nem de 1000 produtos com 4 e 5 ficheiros porque o ficheiro ZIP ultrapassava 500Mb e o tempo de execução seria demasiado longo, e não apresenta relevância para o resultado.

prod.\fich	1 file		2 file		3 file		4 file		5 file	
	s/zip	c/zip	s/zip	c/zip	s/zip	c/zip	s/zip	c/zip	s/zip	c/zip
1	0.1s	1.6s	0.2s	1.7s	0.2s	1.7s	0.4s	1.8s	0.7s	2.0s
10	0.7s	1.7s	1.1s	2.0s	1.4s	2.3s	3.8s	4.2s	6.2s	7.2s
100	4.0s	3.7s	7.6s	6.0s	11.2s	8.0s	34.9s	50.8s	58.6s	76.2s
1000	36.6s	26.8s	72.6s	50.0s	108.6s	93.3s	345.6s	-	582.6s	-
10000	376s	-	736.6s	-	1096.6s	-	3466.6s	-	5836.6s	-

Tabela 16 - Diferença performance ZIP e Simple XLS

Os resultados obtidos permitem observar que para uma quantidade reduzida de ficheiros, a importação através do ficheiro ZIP não compensa em termos de performance. Os melhores resultados de importação verificam-se apenas nos casos em que se importa 100 ou 1000 produtos com imagens associadas. Nos restantes casos a importação tradicional revelou-se mais eficiente.

Isto acontece porque com poucos ficheiros existe uma maior necessidade de processamento para os extrair e processar. Por outro lado, com muitos ficheiros e ficheiros muito pesados, como por exemplo PDF's, o upload e a extração do ficheiro zip não é compensado no processamento de ficheiro. Quando o ficheiro zip tem muitos ficheiros e não ocupa muito espaço o ganho acontece no processamento de ficheiros, tornando-se ligeiramente mais rápido.

Através destes resultados chegamos à conclusão que a rapidez de importação é muito semelhante. A maior vantagem reside na facilidade da importação, porque os ficheiros xls e todos os ficheiros a associar estão no mesmo sítio e são importados de uma só vez sem ser necessário executar vários passos para concluir a importa.

Durante a elaboração do portal existiram várias abordagens para melhorar a performance das operações. Vários foram os procedimentos de código que tivemos que refazer para melhorar a eficiência e eficácia da tarefa. A adaptação a diferentes bibliotecas python também foi uma das partes essenciais para escolher as ferramentas mais adequadas para o tratamento de imagens, pdf's e ficheiros xls.

De forma genérica é possível afirmar que as tecnologias e as ferramentas foram adequadamente escolhidas para todas as importações/exportações. Os tempos de execução são relativamente baixos o que permite reduzir o tempo que o utilizador dispensa a importar os produtos para o portal.

Comparação com os outros sistemas

Depois da elaboração de todo o projeto é possível comparar as ferramentas desenvolvidas com as ferramentas que analisei no ponto 2.5 Análise comparativa. Através desta comparação é possível rapidamente obter os pontos positivos da ferramenta que desenvolvi. Na Tabela 17 está a comparação de importação de produtos para o portal e na Tabela 18 está a comparação de importação dos ficheiros para o portal.

Descrição\Sistemas	Alibaba	eBay	Amazon	Nova
Tipo de importação	Xls	txt	txt e macro	zip, xls ouxlsx
Quantidade de produtos a importar	+/- 900 produtos	50-5000 produtos	50000 produtos	Limite do ficheiro
Descrição do excel	Sim	Sim	Sim	Sim
Dropdown	Sim	Não	Não	Sim
Associar imagens	Não	Sim, uma por link	Sim, por link	Sim
Associar pdf	Não	Não	Não	Sim
Detalhe de erros	No novo excel		Lista de erros e no excel	Lista de erros
Importação em mais do que uma língua	Não	Não	Não	Sim
Tamanho máximo do ficheiro	300 Kb	Não tem	60 Mb	30 Mb
Facilidade de importação(1-difícil até 4-fácil)	3	2	2	4
Línguas do ficheiro	Inglês	Inglês	Inglês	7 línguas
Exportação xls vazio	Árvore de categorias	Sem categorias	Diferentes downloads do xls	Árvore de categorias
Exportação com dados	Não	Não	Não	Sim

Tabela 17 - Comparação diferentes sistemas de importação de produtos

Descrição\Sistema	Alibaba	Dropbox	Novo
Tipo de ficheiros	jpg, jpeg, png	Todos	pdf, jpg, jpeg, png
Tamanho máximo	3Mb	10Gb	3 Mb
Nº de ficheiros a importar	6	Sem limite	Sem limite
Upload feedback	Geral	Em cada ficheiro	Geral
Apresentação do erro	Sim, com detalhe	Sim	Sim, com detalhe
Validação da extensão	Sim	Não	Sim
Validação do ficheiro	Sim	Não	Sim

Tabela 18 - Comparação de diferentes sistemas de importação de ficheiros

6. Conclusão

No meu percurso formativo tive oportunidade de contactar e aprender várias linguagens de programação, umas de baixo nível em que a linguagem se aproxima da linguagem máquina, outras muito próximas da linguagem humana, ou seja, denominadas por linguagens de alto nível. Estive envolvido em projetos de diferentes áreas, desde programação à gestão, permitindo-me desenvolver capacidades que vão muito além da lógica habitual de um programador.

O começo deste projeto que envolveu áreas totalmente novas para mim e com tecnologias que nunca tinha utilizado anteriormente constitui um desafio que me fez pôr à prova as bases adquiridas ao longo do meu percurso académico e que se mostraram uma mais valia na rápida adaptação a um contexto tecnológico diferente. Neste projeto passei por várias etapas desde do início de construção de simples tabelas de base de dados até à tradução de todo o portal para a minha língua materna. Concluída a parte funcional do projeto foi necessária a revisão de todas as especificações tendo em vista a otimização ao máximo da performance. Para alguns casos foi necessário refazer o código por completo, garantindo dessa forma a rapidez na execução das tarefas desenhadas e programadas. Isto permitiu-me aprender que muitas vezes é necessário rever e refazer o nosso trabalho sempre com olhar crítico de forma a encontrar a solução ótima. Aprendi ainda a grande importância do planeamento no desenvolvimento de uma aplicação desta dimensão. As tecnologias que foram escolhidas inicialmente conseguiram provar a estabilidade e os métodos de implementação. Essas tecnologias deram-me a conhecer ferramentas ideais para a construção de uma página web suportada por uma programação simples e flexível. A framework Django revelou bons resultados em termos de utilização e facilidade de implementação. A documentação disponibilizada teve um papel essencial na construção de um portal relativamente complexo. A linguagem Python com que o Django trabalha deu bons resultados em termos de bibliotecas disponíveis para tratamento de operações de forma simples e rápida. A maioria das bibliotecas era ainda acompanhada por uma boa documentação, e de muitas fontes de informação alternativas. A base de dados Postgresql é facilmente integrada e gerida no projeto Django tendo mostrado ao longo de grandes importações de produtos que suporta uma grande dimensão de dados mantendo sempre a rapidez.

Ao longo de estudo da importação de produtos para o portal também concluí que o processo de importação em si reduz o tempo e o esforço utilizado pelo utilizador. Isso prova que o portal é uma ferramenta eficiente, prática e avançada que permite efetuar importações minimizando o tempo gasto pelo utilizador final. A motivação principal no desenvolvimento desta tese prendeu-se com o desafio de construir uma ferramenta simples e fácil de utilizar que poupasse tempo ao utilizador final que pretenda inserir produtos no portal. Foi ainda igualmente motivador construir uma plataforma que permite poupar recursos no que diz respeito à tradução de produtos para todas as línguas e de forma automática e ainda permitir uma melhor divulgação dos produtos através de várias fontes.

6.1 Trabalhos futuros

Para os trabalhos futuros poderão existir várias abordagens para melhorar a comunicação com o próprio portal, permitir uma publicidade mais alargada de produtos, otimizar a performance etc. Apresento ainda alguns exemplos de melhoramentos que já se encontram planeados:

- Ligar os utilizadores do portal pela API, para que as adições/modificações sejam feitas da forma instantânea pelo utilizador. Neste momento o portal ainda está na fase de alargamento e junção das pequenas e médias empresas, sendo este o motivo de ainda não ser usada API de ligação;
- Enviar emails para os utilizadores do portal e para os utilizadores que foram subscritos com as informações de novas empresas que entrar com os seus produtos no portal;
- Suportar importação de produtos através de diferentes tipos de ficheiros, como csv, txt, etc.
- Exportação da lista de produtos duma empresa para um ficheiro PDF, para que o utilizador possa utilizá-lo, por exemplo num contexto comercial;
- Construir um sistema que possa exportar todas as imagens dos produtos do site do utilizador.

6.2 Aprendizagem

Ao longo deste trabalho foi necessário aprender várias tecnologias e utiliza-las de forma correta, tendo a maioria sido totalmente novas para mim. Neste seguimento, consigo afirmar que a minha adaptação a esta nova realidade tecnológica foi rápida e eficaz, não tendo demonstrado grandes dificuldades a este nível. Durante o projeto foi ainda necessário aprender o funcionamento e a lógica de diferentes tipos de ficheiros de forma a melhorar a performance.

Foram também melhorados os conhecimentos já adquiridos na universidade, como por exemplo o estudo e o desenho de base de dados sempre com o objetivo de otimizar a performance. A nível de linguagens de programação utilizadas, algumas delas já conhecidas por mim, foram as seguintes:

- HTML
- JavaScript
- Django
- Python
- Postgresql
- MySQL
- Web-services XML-RPC
- XML
- JSON
- Diversas API's

7. Bibliografia

- [1] Alibaba.com, "Manufacturers, Suppliers, Exporters, Importers from the world's largest online B2B marketplace.," 1999-2015. [Online]. Available: <http://www.alibaba.com/>. [Acedido em 14 Fevereiro 2014].
- [2] Alibaba.com, "Alibaba Manufacturer Directory - Export/Import excel file," 1999-2015. [Online]. Available: <http://us.productposting.alibaba.com/product/upload/introduce.htm>. [Acedido em 14 Fevereiro 2014].
- [3] Alibaba.com, "Alibaba Manufacturer Directory - Import file data," 1999-2015. [Online]. Available: <http://us.my.alibaba.com/photobank/photobank.htm>. [Acedido em 14 Fevereiro 2014].
- [4] eBay Inc., "Electronics, Cars, Fashion, Collectibles, Coupons and More Online Shopping," 1995-2015. [Online]. Available: <http://www.ebay.com/>. [Acedido em 14 Fevereiro 2014].
- [5] eBay Inc, "Import xls file," Fevereiro 2015. [Online]. Available: http://pics.ebay.com/aw/pics/pdf/us/file_exchange/SMP_FMC_User_Guide.pdf. [Acedido em 14 Maio 2015].
- [6] Amazon.com, "Amazon.com: Online Shopping for Electronics, Apparel, Computers, Books, DVDs & more," 1996-2015. [Online]. Available: <http://www.amazon.com/>. [Acedido em 14 Fevereiro 2014].
- [7] Amazon.com, Inc., "Amazon.com Help: Building Your Spreadsheet," 1996-2015. [Online]. Available: <http://www.amazon.com/gp/help/customer/display.html?nodeId=200203300>. [Acedido em 21 Maio 2015].
- [8] Amazon.com, Inc., "Amazon.com Help: Inventory File Templates," 1996-2015. [Online]. Available: http://www.amazon.com/gp/help/customer/display.html?nodeId=200186090#templates_tables. [Acedido em 21 Maio 2015].
- [9] Amazon.com, Inc, "mp_feed_erros.pdf," [Online]. Available: https://images-na.ssl-images-amazon.com/images/G/01/rainier/help/mp_feed_errors.pdf. [Acedido em 21 Maio 2015].
- [10] Amazon.com, Inc, "Amazon.com Help: Enabling Macros in Excel," 1996-2015. [Online]. Available: <http://www.amazon.com/gp/help/customer/display.html?nodeId=200411020>. [Acedido em 21 Maio 2015].
- [11] Amazon.com, Inc., "Amazon.com Help: Using Template Macros," 1996-2015. [Online]. Available: <http://www.amazon.com/gp/help/customer/display.html?nodeId=200431480>.

[Acedido em 21 Maio 2015].

- [12] Amazon.com, Inc., “Amazon.com Help: Review File Status and History,” 1996-2015. [Online]. Available: http://www.amazon.com/gp/help/customer/display.html/ref=help_search_1-5?ie=UTF8&nodeId=200828140&qid=1432196673&sr=1-5. [Acedido em 21 Maio 2015].
- [13] Dropbox Inc., “Dropbox,” [Online]. Available: <https://www.dropbox.com/>. [Acedido em 23 Maio 2015].
- [14] Dropbox Inc., “Dropbox Help Center,” [Online]. Available: https://www.dropbox.com/help/5?path=space_and_storage. [Acedido em 23 Maio 2015].
- [15] MAGENTO, INC., “eCommerce Software & eCommerce Platform Solutions,” 2015. [Online]. Available: <http://magento.com/>. [Acedido em 15 Maio 2015].
- [16] MAGENTO, INC., “eBay Enterprise is #1,” 2015. [Online]. Available: <http://magento.com/blog/magento-news/ebay-enterprise-1-ecommerce-platform-provider-ir-top-500>. [Acedido em 26 Maio 2015].
- [17] Wikipedia Foundation, Inc., “Comércio Eletrônico,” [Online]. Available: http://pt.wikipedia.org/wiki/Com%C3%A9rcio_eletr%C3%B4nico. [Acedido em 26 Maio 2015].
- [18] MailChimp, “Send Better Email,” 2001-2015. [Online]. Available: <http://mailchimp.com/>. [Acedido em 11 Maio 2015].
- [19] Google, [Online]. Available: http://translate.google.pt/about/intl/pt-PT_ALL/contribute.html. [Acedido em 26 Maio 2015].
- [20] Python Software Foundation, “Goslate: Free Google Translate API,” 2013. [Online]. Available: <http://pythonhosted.org/goslate/>. [Acedido em 09 Maio 2015].
- [21] Django Software Foundation, “The Web framework for perfectionists with deadlines,” 2005-2015. [Online]. Available: <https://www.djangoproject.com/>. [Acedido em 14 Fevereiro 2014].
- [22] Python Software Foundation, 2001-2015. [Online]. Available: <https://www.python.org/>. [Acedido em 14 Fevereiro 2014].
- [23] Adrian Holovaty, Jacob Kaplan-Moss, “Chapter 5: Models,” [Online]. Available: <http://www.djangobook.com/en/2.0/chapter05.html#the-mtv-or-mvc-development-pattern>. [Acedido em 24 Maio 2015].
- [24] Wikipedia Foundation, Inc., “Model View Controller,” 2015. [Online]. Available: <http://en.wikipedia.org/wiki/Model%E2%80%93view%E2%80%93controller>. [Acedido em 20

Maio 2015].

- [25] Microsoft, “Python web app with Django,” 2015. [Online]. Available: <http://azure.microsoft.com/pt-pt/documentation/articles/virtual-machines-python-django-web-app-linux/>. [Acedido em 24 Maio 2015].
- [26] The PostgreSQL Global Development Group, “PostgreSQL: The world's most advanced open source database,” 1996-2015. [Online]. Available: <http://www.postgresql.org>. [Acedido em 14 Fevereiro 2014].
- [27] TM, “PostgreSQL vs MySQL - Habrahabr,” [Online]. Available: <http://habrahabr.ru/company/mailru/blog/248845/>. [Acedido em 12 Maio 2015].
- [28] Wikipedia Foundation, Inc., “PostgreSQL,” 2015. [Online]. Available: <https://ru.wikipedia.org/wiki/PostgreSQL>. [Acedido em 15 Maio 2015].
- [29] GS1 AISBL, “GPC Browser | Global Product Classification | Standards,” 2014. [Online]. Available: <http://www.gs1.org/1/productssolutions/gdsn/gpc/browser/>. [Acedido em Abril 2014].
- [30] J. McNamara, “Creating Excel files with Python and XlsxWriter,” 2013-2015. [Online]. Available: <https://xlsxwriter.readthedocs.org/>. [Acedido em 18 Maio 2015].
- [31] Django Software Foundation, “File Uploads,” 2005-2015. [Online]. Available: <https://docs.djangoproject.com/en/1.5/topics/http/file-uploads/>. [Acedido em 18 Maio 2015].
- [32] Simplistix Ltd., “The xlrd Module - Read file,” Março 2012. [Online]. Available: <https://secure.simplistix.co.uk/svn/xlrd/trunk/xlrd/doc/xlrd.html?p=4966>. [Acedido em 15 Maio 2015].
- [33] Django Software Foundation, “Database transactions,” 2005-2015. [Online]. Available: <https://docs.djangoproject.com/en/1.8/topics/db/transactions/>. [Acedido em 12 Maio 2015].
- [34] Amazon Web Services, “AWS | Amazon SimpleDB – Simple Database Service,” 2015. [Online]. Available: <http://aws.amazon.com/simplydb/>. [Acedido em 18 Maio 2015].
- [35] Python Software Foundation, “django-watson 1.1.8,” 1990-2014. [Online]. Available: <https://pypi.python.org/pypi/django-watson>. [Acedido em 18 Maio 2015].
- [36] Django Project, “QuerySet API Bulk Insert,” [Online]. Available: <https://docs.djangoproject.com/en/1.8/ref/models/querysets/#bulk-create>. [Acedido em 18 Maio 2015].

- [37] Django Software Foundation, “Managers - connection cursor,” 2005-2015. [Online]. Available: <https://docs.djangoproject.com/en/1.8/topics/db/managers/#adding-extra-manager-methods>. [Acedido em 18 Maio 2015].
- [38] Y. Shinyama, “Programming with PDFMiner,” 2014. [Online]. Available: <http://www.unixuser.org/~euske/python/pdfminer/programming.html>. [Acedido em 18 Maio 2015].
- [39] effbot.org, “Python Imaging Library Handbook,” [Online]. Available: <http://effbot.org/imagingbook/pil-index.htm>. [Acedido em 18 Maio 2015].
- [40] Python Software Foundation, “UUID objects according to RFC 4122,” 1990-2015. [Online]. Available: <https://docs.python.org/2/library/uuid.html>. [Acedido em 18 Maio 2015].
- [41] Wikipedia Foundation, Inc., “Zip (file format),” 2015. [Online]. Available: [en.wikipedia.org/wiki/Zip_\(file_format\)](http://en.wikipedia.org/wiki/Zip_(file_format)). [Acedido em 17 Maio 2015].
- [42] Python Software Foundation, “zipfile - work with ZIP archives,” 1990-2015. [Online]. Available: <https://docs.python.org/2/library/zipfile.html>. [Acedido em 20 Maio 2015].
- [43] MAGENTO, INC., “Magento API - XMLRPC,” 2015. [Online]. Available: <http://www.magentocommerce.com/api/soap/introduction.html#Introduction-XMLRPC>. [Acedido em 18 Maio 2015].
- [44] MAGENTO, INC., “Magento API - create product,” 2015. [Online]. Available: http://www.magentocommerce.com/api/soap/catalog/catalogProduct/catalog_product.create.html. [Acedido em 18 Maio 2015].
- [45] MAGENTO, INC., “Magento API - create media,” 2015. [Online]. Available: http://www.magentocommerce.com/api/soap/catalog/catalogProductAttributeMedia/catalog_product_attribute_media.create.html. [Acedido em 18 Maio 2015].
- [46] Oracle Corporation, “MySQL,” 2015. [Online]. Available: <https://www.mysql.com/>. [Acedido em 15 Maio 2015].
- [47] Python Software Foundation, “base64 - RFC 3548,” 1990-2015. [Online]. Available: <https://docs.python.org/2/library/base64.html>. [Acedido em 17 Maio 2015].
- [48] Python Software Foundation, “19.8 xml.dom,” 1990-2015. [Online]. Available: <https://docs.python.org/2/library/xml.dom.html>. [Acedido em 17 Maio 2015].
- [49] MailChimp, “MailChimp API,” 2001-2012. [Online]. Available: <https://apidocs.mailchimp.com/api/>. [Acedido em 19 Maio 2015].

- [50] Google Project Hosting, "Python wrapper for the MailChimp API," [Online]. Available: <https://code.google.com/p/chimpy/>. [Acedido em 19 Maio 2015].
- [51] Artima, Inc., "The DCI Architecture: A New Vision of Object-Oriented Programming," 1996-2015. [Online]. Available: http://www.artima.com/articles/dci_vision.html. [Acedido em 20 Maio 2015].
- [52] The PostgreSQL Global Development Group, "PostgreSQL: About," 1996-2015. [Online]. Available: <http://www.postgresql.org/about/>. [Acedido em 15 Maio 2015].
- [53] I. Paul Lindner, "Registration of a new MIME Content-Type/Subtype - application/zip," 20 Julho 1993. [Online]. Available: <http://www.iana.org/assignments/media-types/application/zip>. [Acedido em 17 Maio 2015].
- [54] D. Greenfeld e A. Roy, "Two Scoops of Django," 2013, pp. 17-20.
- [55] D. Greenfeld e A. Roy, "Two Scoops of Django," 2013, pp. 218-219.